ANALYSIS OF A CONTROLLABLE M/M/2 QUEUEING SYSTEM OPERATING UNDER THE TRIADIC (0,K,N,M) POLICY

BY

HAHN-KYOU RHEE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1989

ACKNOWLEDGEMENTS

The author would like to express his sincere appreciation to Dr. Boghos D. Sivazlian, chairman of the author's supervisory committee, for all his invaluable advice, guidance and encouragement throughout this research.

Special appreciation is extended to Dr. John G. Saw, Dr. Suleyman Tufekci, Dr. Chung-Yee Lee and Dr. Sencer Yeralan for serving on the author's supervisory committee and for their suggestions, advice and guidance.

Appreciation is also extended to the Department of Industrial and Systems Engineering at the University of Florida for providing financial support

Finally, the author wishes to express the most heartfelt appreciation to his parents and sisters for their support and encouragement and to his wife for her continuous support, patience and encouragement.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ANALYSIS OF A CONTROLLABLE M/M/2 QUEUEING SYSTEM OPERATING UNDER THE TRIADIC (0,K,N,M) POLICY

BY

HAHN-KYOU RHEE

May, 1989

Chairman: Boghos D. Sivazlian
Major Department: Industrial and Systems Engineering

We consider an M/M/2 queueing system with removable service stations operating under steady-state over an infinite time horizon. We assume that the number of operating service stations can be adjusted at customers' arrival and service completion epochs depending on the number of customers in the system.

First, the steady-state behavior of the system is studied as a birth and death process. The joint probability mass function of the number of customers in the system and the number of operating service stations is obtained as well as other system characteristics which might be incorporated in a total expected cost function to optimize the operation of the system.

Next, the busy period is analyzed using the properties of the classical gambler's ruin problem. Fundamental information of system

behavior such as the probability mass function of the number of activations or removals of service stations during the busy period and the Laplace transform of the probability density function of the busy period are derived. As a special case, a number of existing results such as the Laplace transforms of the probability density functions of the busy periods for the ordinary M/M/2 system, the M/M/1 system operation under the N policy and the ordinary M/M/1 system, are recovered. Alternative methodologies are also provided to find certain system characteristics of particular use in a cost model.

Finally, a total expected cost function per unit time is constructed to obtain the optimal customer levels dictated by the three decision variables at which adjustments of service stations are made. A sensitivity analysis is also performed for changes of the optimal customer levels along with changes in specific cost elements.

CHAPTER 1

INTRODUCTION

1.1 General Description of Controllable Queueing Models

Almost all the fields of modern society have faced the problem of providing huge demands for simultaneous service. This reality makes waiting lines an unavoidable component of modern life style leading to accelerated research activities in waiting line systems. The general trend of research has been to create mathematical models appropriately based on real waiting line situations and to derive important system characteristics using these models.

In the early stages of queueing theory, the basic assumptions for the creation of mathematical models involved fixed system parameters. In other words, each system parameter could not be allowed to change over time or system state. The basic assumptions were critical factors not only to describe in a simple fashion the real situations, but to get desired results easily. However, although the results obtained could provide information to predict real system behavior, there seems to be a large gap between assumed models and real situations because the basic assumptions might be far from realistic. This motivates a change of research direction by considering models which could afford relaxations of those assumptions.

The controllable queueing models which allow the system parameters to be changed over time or system state have been a major part of

research in queueing theory over the last two decades. Numerous different models which could be considered as control models have been introduced in the literature. There could be a number of different criteria to classify the introduced models. However, Crabil, Gross and Magazine[9] and J.Teghem Jr.[25] distinguish four categories for these models depending on which system parameter can be changed over time. These categories are 'control of the number of servers or service stations', 'control of the service rate', 'control of the admission of the arrival customers' and 'control of the waiting line discipline'.

Here we are concerned with the models related to the 'control of the number of service stations'; the so called vacation models. The main purpose of the vacation models is to try to utilize the server's time. There could be partial or complete removals of service stations based on a predetermined operating policy. Thus, the criteria for activations and removals of the service stations and the epochs at which those adjustments should occur are the most important assumptions of the operating policy. Review points or decision points are defined on the time axis as those epochs at which the decisions are made on whether the adjustments of the service stations should occur or not. According to the definition of review points, the vacation models can be classified into three categories. The model is called the periodic review model if the time intervals between any two successive review points are the same and determinations of the review points are independent of the system states. If the review points are stochastically determined depending on the system states, then the model is called the stochastic review model. In this model, the review points are not necessarily equally spaced in time. The last type is the continuous review model in which the

adjustments can be made continuously over time. The customers' arrival and service completion epochs are one of the possible choices as review points in the area of vacation models. The criteria for adjustments of service stations are closely related to the decision variables. Based on the activation criteria of removed service station, three different types of operating policy have been introduced. The first one is called the N policy in which the decision variables for activations of service stations are the number of customers in the system (see Yadin and Naor[27]). Thus, the removed service station can only be activated at a time when the number of customers in the system reaches a specific level. If the decision variables are workloads which are the total amounts of existing customers' work in the system or server's backlog, D, then the operating policy is called the D policy (see Balachandran and Tijms[1] and Boxma[6]). Finally, an operating policy is called the T policy if service stations only activate T time units after the end of the last busy period (see Heyman[12]). Furthermore, according to the removal criteria of activating service stations, one can have the exhaustive policy or the hysteretic policy. If the activated service stations can only be removed at a time when the service stations become idle, then one has the exhaustive policy. The other case is the hysteretic policy in which service stations are removed even when they do not become idle.

The general approach used in the controllable queueing problems is to decide on the optimal operating policy. This can be done based on minimizing the total expected cost involved in system operation. Usually, the cost elements consist of holding cost, operating cost, and switching cost. The holding cost depends on the number of customers in

the system. The operating cost is related to the number of operating service stations. The adjustments of service stations contribute to the switching cost.

1.2 Literature Review

The paper of Moder and Phillips Jr.[19] is a typical example which shows the general research pattern dealing with the control of the number of service stations in the beginning years of controllable queueing models. During this time, the prime interest of researchers was to derive the only system characteristics which could be related to cost elements under the assumed operating rules. Even though the reasonable operating rules were well developed, optimality of the assumed operating rules was not discussed.

More recently the interests of researchers in the area of removable servers or service stations have changed. The developments of realistic operating policies such as the N policy, the D policy and the T policy, and the determinations of the optimal decision variables have been of major interest. However, we concentrate on those papers dealing with the N policy. At this stage, it should be pointed out that most of the work in the area of N policy are concerned with models with a single removable service station. Only a few papers devote themselves to the multi-removable service stations operating under the concept of the N policy.

The next section is divided into two parts according to the number of removable service stations studied in the literature. The first part reviews those papers which treat the single removable service station operating under an N policy. The second part deals with those papers

treating the multi-removable service station models.

1.2.1 Single Removable Service Station Operating under an N policy

The concept of an N policy in queueing theory was first introduced by Yadin and Naor[27]. They consider an M/G/1 system in which the server is removed at the end of a busy period. Then, the server is activated when the number of customers in the system reaches R for the first time. They derive the expected number of customers in the system for the removable M/G/1 system. They also derive the optimal decision variable R^* which minimizes the total expected cost per unit time including the holding costs, the switching costs and certain negative costs (profits) which are the results of utilization of server's idle time.

Heyman[11] also considers an M/G/1 queueing system with removable service stations. His cost elements consist of the operating cost, the holding cost and the switching cost. He investigates the optimal behaviors of the system under the average cost and the discounted cost criterion over an infinite time horizon. For the average cost criterion, he derives explicitly the optimal customer level at which the service station should be opened. For the discounted cost criterion, he proves the same general properties of the optimal operating N-policy as those of the average cost criterion. However, because of the difficulties in finding the optimal customer level explicitly, he provides an algorithm to determine it indirectly.

Sobel[22] studies a GI/G/1 removable service station with an N policy using the average cost criterion over an infinite time horizon. He shows that Heyman's results on the form of the optimal operating

policy are valid for his model.

Kimura[14] applies diffusion approximation method to Heyman[11]'s discounted cost model. He uses diffusion processes to approximate the number of customers in the system by a continuous variable. He shows numerically that his approach can generate the Heyman's results under heavy traffic condition.

Several interesting extensions of the above models can also be found in the literature. But we only introduce here a brief description of these models. A more detailed review appears in the paper of Teghem, Jr.[25].

Some researchers have studied the single removable server queueing system operating under an N policy for a finite system capacity. For example, one can cite Hersh and Brosh (see Teghem, Jr.[25]) for the M/M/1 system, and Lolis and Teghem (see Teghem, Jr.[25]) for the M/G/1 system. In these types of models, the holding costs are usually replaced by the penalty costs incurred for lost customers.

Others have extended the basic models using priority queueing system. Bell[2,3] considers an M/G/1 priority queue with removable server using an infinite horizon average cost as the criterion. He assumes several priority classes of customers, each class of customer having the different holding cost and the same service time distribution. Tijms (see Teghem, Jr.[25]) studies an M/G/1 non-preemptive priority queueing system with a removable server and two priority classes of customers. He assumes that each class of customers has a different holding cost and a different service time distribution.

1.2.2 Multi-removable Service Stations

Magazine[16] considers an M/M/s queueing system with removable service stations under periodic review. The number of customers in the system is observed at equally spaced review points in time. His decision variables are the number of servers. Decisions may be made at each review point. If the observed number of customers in the system at each review point falls between two predetermined values, then a decision is made on how many servers should be activated or removed. If it falls outside the range of two values, no action is taken. Using dynamic programming, he shows the existence of a control limit policy for the average cost criterion over a finite and an infinite time horizon. In other words, there exists a unique number of customers in the system, say j^*, such that in order to activate the service station the observed number of customers in the system should be greater than j^*. He also proves the control limit policy is optimal for the infinite horizon model and the average cost criterion.

Magazine[17] extends his previous model under assumption of a finite system capacity. He also derives the same conclusions as those of his previous model.

Huang, Brumelle, Sawaki and Vertinsky[13] consider a similar model to the one outlined by Magazine[17], with some modifications. They assume infinite system capacity and generalized holding costs under a discounted cost criterion over a finite time horizon. They permit different holding costs which are functions of the number of customers in the system at the end of each period. They also assume that the holding cost functions are convex. They conclude that the optimal operating policy has the form of a control limit policy.

Szarkowicz and Knowles[23] investigate the model studied by Huang et al.[13] under less restrictive assumptions. But they derive the same conclusions as those of Huang et al.

In the paper by McGill[18], an M/M/s and a GI/M/s queueing system are investigated under stochastic review. He assumes that the time intervals between any two successive review points are random variables which depend on the system states. Like Magazine[17], his decision variables are also the number of servers to be activated or removed based on the number of customers in the system. He develops a framework for determination of the optimal decision variables under total expected discounted and undiscounted costs criteria for the M/M/s system. He derives the explicit optimal control policy for the case of no switching cost.

Bell[4] treats an M/M/s model studied by McGill. The number of servers working can be adjusted at customers' arrival or service completion epochs. He defines an efficient policy as an operating policy which never allows more active service stations than the number of customers presented in the system, otherwise an operating policy is inefficient. He shows numerically that the optimal operating policy may not necessarily be an efficient policy.

Winston[26] examines a removable M/M/s system in which the customers' arrival rates depend on the current system state. His review points are also the customers' arrival and service completion epochs, and his decision variables are the number of servers. Unlike other models, he assumes that the removal or the activation of service stations incurs no costs. But he generalizes the operating costs and the holding costs. His holding costs are a function of the number of

customers in the system. His operating costs are also a function of the number of active service stations. He derives optimal conditions that ensure that the number of servers in operation is a non-decreasing function of the number of customers in the system.

More recently Bell[5] tries to apply the concept of an N policy to the multi-removable service station case. He considers an M/M/2 queueing system where the customers' arrival and service completion epochs are the review points. He defines the (S_0,S_1,R_1,R_2) policy denoting the number of customers in the system when the number of opening servers should be adjusted downward to 0,1 and upward to 1,2, respectively. He allows the possibility that the number of opening service stations is never adjusted downward to 0 or 1, denoting by $S_0 = -1$, or $S_1 = -1$. Without deriving the system characteristics explicitly, he shows that the optimal operating policy is an efficient policy if the operating costs are sufficiently high. If the operating costs are sufficiently close to zero compared to the holding costs and switching costs, an inefficient policy is optimal. He also proves general properties of the optimal operating policy. For the efficient policy, the optimal policy has the form $S_0 = 0$, $1 < S_1 < R_2$. For the inefficient policy, the optimal operating policy may be one of following cases: (1) leave both servers opened at all times, denoted by $S_0 = S_1 = -1$, (2) leave at least one server opened at all times, denoted by $S_0 = -1$, $S_1 \geq 0$, $S_1 < R_2$, or (3) always close both servers, denoted by $S_0 = S_1 = 0$, $R_2 \geq R_1$.

Even though most researchers have concentrated on deriving the optimal operating policy rather than analyzing the system behavior in the area of multi-removable service stations, no one has attempted to

derive it explicitly because of the complexity of the problem.

1.3 Problem Statement

In this research, we consider an M/M/2 queueing system with removable service stations under steady-state over an infinite time horizon. The review points of the system are the customers' arrival and service completion epochs. The decisions on the adjustment of the service stations are made based on the number of customers in the system. Thus, the operating policy of our model should be considered a type of the N policy. The decision variables of our model are several levels of customers in the system at which the adjustments of the service stations should occur.

1.3.1 Operating Policy of the System

Consider an M/M/2 steady-state queueing process. The customers' interarrival times are identically, independently and exponentially distributed. The customers' service times at both service stations are also identically, independently and exponentially distributed. The customers' interarrival times and the customers' service times are independently distributed. The special feature of the operating policy of this queueing system is that the number of operating service stations can always be adjusted based on the number of customers present in the system. Thus, the number of customers in the system is monitored at every new customer's arrival and service completion epochs. Whenever there are no customers in the system, both service stations remain inoperative temporarily and may not be reactivated until certain conditions are satisfied. Suppose both service stations are removed

from the system. When the number of customers waiting for service reaches a specified level, denoted by N which is a positive integer, one of the two service stations removed from the system is activated immediately. At a later time when the number of customers in the system increases to another specified level, say M where $0 < N < M$, then the remaining service station removed from the system is also activated immediately. However, if the number of customers in the system decreases to K where $1 \leq K \leq N$ while both service stations are operating, the service station just finishing a service will be removed at that instant. Furthermore, if the number of customers in the system reaches zero while one service station is operating, that service station should also be removed until the above reactivating conditions are satisfied.

We denote the operating policy of our model as the (0,K,N,M) policy. This operating policy is an efficient policy according to Bell[4]'s definition because the number of active service stations is always less than the number of customers present in the system. Furthermore, no service stations are allowed to operate when the system becomes empty. Thus, this operating policy can be considered as a type of the exhaustive policy.

1.3.2 Special Cases of the Model

We assumed that the number of operating service station could be adjusted based on the number of customers in the system. Hence, unknown values of the decision variables K, N and M are crucial criteria to determine the system operation. This implies that each value of the decision variables specifies a different operating policy of the system.

Now, we consider two typical special cases of the (0,K,N,M) operating policy by assigning specific values to the decision variables.

Case 1 : Suppose that K=1, N=1 and M=2, denoted by (0,1,1,2) policy. Then, whenever there is a customer in the system, the number of operating service stations is one. If the number of customers in the system is greater than or equal to two, both service stations are operating simultaneously. When the system becomes empty, no service stations are operating. Hence, it is easy to see that the M/M/2 queueing system operating under the (0,1,1,2) policy is equivalent to the ordinary M/M/2 queueing system.

Case 2 : Suppose that $M=\infty$, denoted by $(0,K,N,\infty)$ policy. One of the two service stations removed from the system starts operating when the number of customers in the system reaches N. Since $M=\infty$, the remaining service station removed from the system cannot start operating at all. Hence, the operating service station can only be removed when there are no customers in the system. In this case, the behavior of the system is independent of the value of K. Thus, the behavior of the M/M/2 queueing system operating under the $(0,K,N,\infty)$ policy becomes that of the M/M/1 queueing system operating under the N policy which is studied by Yadin and Naor[27].

We thus conclude that the M/M/2 queueing system operating under the (0,K,N,M) policy is considered as a generalized model of the ordinary M/M/2 queueing system and the M/M/1 queueing system operating under the N policy.

1.3.3 Applicable Examples of the Model

Our model could be applicable to approximate a number of real situations such as banks, post offices, production planning, manufacturing, communication and so on. We elaborate on two examples.

First, we give an example in manufacturing area.

Example 1 : Consider a manufacturing plant where its major tasks are assembly, inspection and rework. Suppose that a large number of workers assemble the same type of expensive items individually at the assembly department. All assembled items are tested at the inspection department. Whenever inspectors find a defective item, they send it to the rework department one by one. Interarrival times of defective items at the rework department are exponentially distributed. Each defective item could be corrected by additional work or could be disassembled to reutilize some components of defective items at the rework department. However, all the defective items arriving at the rework department are kept in special type of containers in order to protect them from further damage. The corrections and the disassembly for defective items can be done effectively by ones who are familiar with the assembling process. Approximately, the correction or the disassembly time for an defective item is exponentially distributed. Workloads at the rework department vary depending on the number of defective items arrived. It is not a good strategy to keep the workers always staying at the rework department, because there could be no defective items to be reworked. Thus, to utilize the reworking workers' times, the rework department should be inoperative under certain conditions. This circumstance requires that an operating rule of the rework department has to be

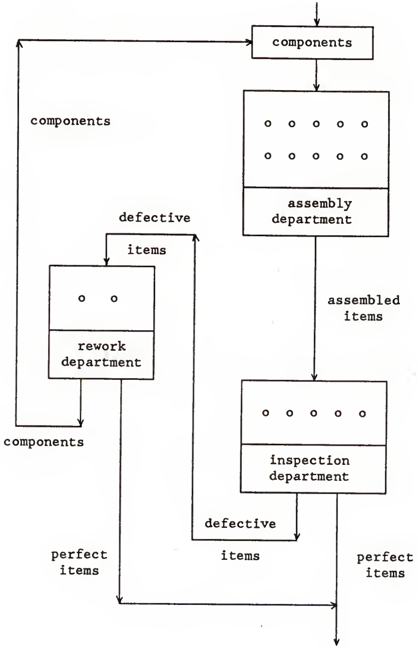

Figure 1.1 The flow diagram of assembling process

related to the number of defective items arrived. In order to analyze this situation, without loss of generality, assume that two of the assembly workers are designated to be part time crew at the rework department. Their tasks are changed alternately depending on where they work, the rework or the assembly. Whenever there are no defective items arriving at the rework department, the rework department is inoperative temporarily. Thus, two designated workers perform the assembling tasks like other assembly workers at the assembly department. If the number of defective items arriving at the rework department reaches N, then one of the two designated workers stops the assembling task and immediately starts the reworking task at the rework department. After that, if the number of arriving defective items increases to M which is greater than N, the other designated worker also joins the rework department. However, if the number of defective items decreases to K where $K \leq N$ while both workers are at the rework department, one of the workers stops the reworking task and returns to the assembly department in order to perform the assembling task. The remaining worker closes the rework department and returns to the assembly department in order to perform the assembling task if there are no defective items at the rework department. Otherwise, the remaining worker performs reworking task for the defective items. The removals and the additions of two designated workers at the assembly department do not significantly change the defective rate of assembling items. This operating policy of the rework department can not only utilize the worker's time, but also assign balancing workloads to each worker. However, this operating policy requires switching costs of the reworking machines and the holding costs of defective items to keep in the containers at the rework department.

The optimal operating policy at the rework department has to be selected based on minimizing the total expected costs.

Another example is related to a communication problem.

Example 2 : Suppose that the main office of a sales company receives important marketing information or messages from several sources. Arrivals of the messages at the main office are assumed to follow a Poisson stream. All messages received are to be sent to a branch office in a foreign country by a commercial satellite system. The company can use two access lines via the satellite to send the messages. The access lines which can be considered as the service stations are available whenever the company wants to send the messages. Transmission time of each message via the satellite is exponentially distributed. But whenever the company uses any available access lines, the company has to pay more charges for a first service time unit and less for subsequent service time units. Thus, a turn-on cost is incurred in the use of the access line which is the excess charge for the first service time unit. On the one hand, it may be economical to turn-on both access lines continuously and to send the messages whenever the main office receives them. Thus, the company can avoid the turn-on costs of access lines even though unnecessary operating costs of the access line are incurred when there are no messages to be sent. On the other hand, the company may experience a certain type of loss in using the access lines selectively because the messages received are kept for a longer time without sending them to the branch office. Therefore, contribution of the turn-on costs, the operating costs of the access lines and the holding costs of the messages received would determine the optimal operating policy of access lines. The optimal decision variables which

minimize the total expected cost involved are the levels of number of messages received at which access lines should be turned-on or turned-off. Therefore, it is reasonable to opt for the following operating policy for the access lines: The main office keeps the messages received without sending them to the branch office until their number reaches a certain level, denoted by N. If that number reaches N, one of the available access lines is used and messages are sent in the order of arrivals until they are all sent to the branch office. If that number reaches a level $M > N$, the other access line starts operating until the number of remaining messages reaches a level $K \leq N$ at which time one of the operating access lines is turned off. One is then interested in the determination of the optimal decision variables (K^*,N^*,M^*) to minimize the total expected cost per unit time.

1.3.4 Basic Assumptions of System Operation

To formulate mathematically the problem, the following assumptions are made.

(1) Customers' arrivals follow a Poisson process with intensity λ. In other words, the random variables A_i where $i=1,2,3,\cdots$ which represent the interarrival times between the (i-1)th and the ith customers joining the system are identically, independently and exponentially distributed with $E[A_i] = 1/\lambda$ for all i. The common probability density function of A_i is denoted by

$$f_A(t) = \lambda e^{-\lambda t} \qquad \lambda \geq 0,\ t \geq 0.$$

(2) The random variables S_i's where $i=1,2,3,\cdots$ which represent the ith customer's service time are identically, independently and exponentially distributed with $E[S_i] = 1/\mu$ for all i. The common

probability density function of S_i is denoted by

$$f_S(t) = \mu e^{-\mu t} \qquad \mu \geq 0, \ t \geq 0.$$

(3) The A_i's and the S_i's are independently distributed.

(4) Both service stations are indistinguishable.

(5) Arriving customers form a single waiting line.

(6) The service stations are not subject to break-downs.

(7) Customers are served in order of their arrivals (First-come, First-served).

(8) Population size and system capacity are infinite.

(9) The system does not allow bulk arrivals, bulk services, balking or reneging.

(10) The system operates under the steady-state conditions which are assumed to exist and which will be determined in the sequel.

1.4 Objectives of Research

Our model is similar to the one outlined by Bell[5], but the following refinements point out the differences between our work and the aforementioned work including Bell's, thus justifying our research.

(1) Rather than be concerned with deriving the optimal operating policy structure explicitly, one of the objectives of this research is to derive the system characteristics which might be incorporated in a total expected cost function without using the explicit form of the steady-state joint probability mass function of the number of customers and the number of operating service stations in the system. Our approach could be applicable to find out the system characteristics of other Markovian queueing problems in which there are some difficulties in obtaining explicitly the steady-state

joint probability mass function of the number of customers and the number of operating service stations in the system.

(2) Another objective of this research is to derive important system characteristics which might not be derivable from the steady-state joint probability mass function of the number of customers and the number of operating service stations in the system. This may involve for example the determination of the probability mass function of the number of activations or removals of service stations during a busy period, the Laplace transforms of the probability density functions of the busy period, the length of time when one service station is operating and the length of time when both service stations are operating simultaneously during the busy period. These system characteristics might be essential to provide fundamental information for analysis of the system behavior. In particular, a number of existing results, namely, the Laplace transforms of the probability density functions of the busy periods for the ordinary M/M/2 system derived by Conolly[8], the M/M/1 system operating under the N policy originated by Yadin and Naor[27] and the ordinary M/M/1 system derived by Conolly[8] or Kleinrock[15] can be recovered from the obtained Laplace transform of the probability density function of the busy period for our model.

(3) Even though the main purpose of removable service stations in queueing system is to utilize the idle service stations, most of the total expected cost functions developed in the area of multi-removable service stations are not concerned with the utilization of the idle service stations as a possible source of cost element.

Furthermore, the same cost structures for each operating service station are assumed regardless of changes in the number of operating service stations. Because of these assumptions, lack of applicability for the total expected cost function to various situations is unavoidable. Hence, total expected cost function per unit time, which is another objective of this research, is constructed by considering the utilization of the idle service stations as a possible source of cost element and assuming different structures for several cost elements based on the number of operating service stations. Thus, this total expected cost function per unit time could be applicable to a variety of situations.

1.5 Overview of the Next Chapters

In Chapter 2, we derive the system characteristics which might be incorporated in a total expected cost function using the explicit form of the steady-state joint probability density function of the number of customers and the number of operating service stations in the system. The advantage of this approach is to obtain certain system characteristics relatively easily from well known formulas. However, although the explicit form of this joint probability mass function is known, nevertheless it is very difficult to find from it the probability mass function of the number of activations or removals of service stations during the busy period, the probability density function of the busy period and so on.

In order to overcome the weakness of the approach used in Chapter 2, we analyze in Chapter 3 the busy period to provide more important

information of the system behavior. The results of the analysis of the busy period provide the fundamental relations to obtain the probability mass function of the number of activations or removals of service stations during the busy period, the Laplace transform of the probability density function of the busy period and so on.

In Chapter 4, we derive the probability mass function of the number of activations or removals of service stations during the busy period using the results of the analysis of the busy period in Chapter 3 and the properties of the classical gambler's ruin problem. The explicit form of this probability mass function plays key role to proceed towards other objectives of this research.

In Chapter 5, we derive the Laplace transforms of the probability density functions of the busy period, the busy cycle, the length of time when one service station is operating and the length of time when both service stations are operating simultaneously during a busy period. We also use the properties of the classical gambler's ruin problem along with the results of Chapter 2 and 3. We recover, as a special case, the Laplace transforms of the probability density functions of the busy period for the ordinary M/M/2 system, the M/M/1 system operating under the N policy and the ordinary M/M/1 system.

In Chapter 6, we develop alternative methodologies to find certain system characteristics that will be used to construct a total expected cost function in Chapter 7 based on the results of Chapter 3, 4 and 5.

In Chapter 7, we construct a total expected cost function per unit time based on the results of the previous chapters. We assume that the cost elements include the cost for the service stations removed from the system, the holding cost for waiting customers, the operating cost for

operating service stations and the switching cost for activations and removals of service stations. We also assume (i) different operating cost for each operating service stations depending on the number of operating service stations similar to the concepts of individual and joint ordering policies in inventory theory and (ii) the switching cost for activations and removals of service stations is not necessarily incurred whenever service stations are activated or removed. The optimal customer levels (K^*, N^*, M^*) at which service stations should be activated and removed are determined numerically in order to minimize the total expected cost per unit time. We also perform a sensitivity analysis for changes of the optimal customer levels along with changes in specific cost elements.

Finally, in Chapter 8, we discuss extension of our methodologies to analyze the busy period for other Markovian queueing models. Topics for future research in the area of controllable M/M/2 queueing system are also suggested.

CHAPTER 2

DERIVATIONS OF SYSTEM CHARACTERISTICS USING THE JOINT PROBABILITY MASS FUNCTION OF THE NUMBER OF CUSTOMERS AND THE NUMBER OF OPERATING SERVICE STATIONS IN THE SYSTEM

2.1 Introduction

The main purpose of this chapter is to derive the important system characteristics for a steady-state M/M/2 queueing system with removable service stations operating under the (0,K,N,M) policy over an infinite time horizon. The methodology used in this chapter requires the explicit form of the steady-state joint probability mass function of the number of customers and number of operating service stations in the system.

The following six system characteristics which might be incorporated in a total expected cost function per unit time are derived explicitly in this chapter.

(1) <u>Steady-State Probability that 'i' Service Stations are Operating and 'j' Customers are in the System, denoted by P(i,j) where i=0,1,2 and j=0,1,2···</u> : The P(i,j)'s are obtained by solving a set of steady-state difference equations recursively.

(2) <u>Probability Generating Function of the Number of Customers in the System, denoted by G(z)</u> : G(z) is derived explicitly using the set of steady-state difference equations.

(3) Expected Number of Customers in the System, denoted by E[J] : From the explicit form of G(z), E[J] is derived by using the property

$$E[J] = \left.\frac{dG(z)}{dz}\right|_{z=1}.$$

(4) Expected Number of Operating Service Stations, denoted by E[O] : Using the well-known Little's formula, E[O] is obtained.

(5) Expected Length of the Idle Period, the Busy Period and the Busy Cycle, denoted by E[I], E[B] and E[C], respectively : Define the idle period, the busy period and the busy cycle as following.

1) Idle period denoted by I ; the length of time when both service stations remain inoperative temporarily.

2) Busy period denoted by B ; the length of time when at least one service station is operating.

3) Busy cycle denoted by C ; the length of time from the beginning of the last idle period to the beginning of the next idle period. In other words, the busy cycle is the sum of the idle period and the busy period.

When both service stations remain inoperative temporarily, no service stations are allowed to operate until the number of customers who are waiting for service in the system reaches N. Hence, the idle period can be represented by the sum of the A_j's where $j=1,2,\cdots,N$ which are the (j-1)th and the jth customers' interarrival times. By the assumptions of system operation, the A_j's are identically and independently distributed with $E[A_j] = 1/\lambda$. Therefore,

$$E[I] = N\ E[A_j].$$

Let $P[O=0]$ be the probability that no service stations are operating. Then, $P[O=0]$ can be computed from

$$P[O=0] = \sum_{j=0}^{N-1} P(0,j).$$

Since

$$P[O=0] = \frac{E[I]}{E[I] + E[B]},$$

then, $E[B]$ is obtained from

$$E[B] = \frac{E[I] - E[I]\,P[O=0]}{P[O=0]}.$$

Thus, $E[C]$ can be derived explicitly from

$$E[C] = E[I] + E[B].$$

(6) <u>Expected Number of Activations (Removals) of Service Stations per Unit Time, denoted by $E[R_a]$ ($E[R_r]$)</u> : Define the random variables R_a and R_r as the number of activations and removals of service stations per unit time, respectively. In any busy cycle, one of the service stations removed from the system should be activated immediately whenever a new customer arrives at the system (i) while N-1 customers are in the system and both service stations remain inoperative temporarily, or (ii) while M-1 customers are in the system and one service station is operating. Since the assumed customers' mean arrival rate is λ, $E[R_a]$ is obtained from

$$E[R_a] = \lambda P(0,N-1) + \lambda P(1,M-1).$$

Similarly, $E[R_r]$ is obtained form

$$E[R_r] = \mu P(1,1) + 2\mu P(2,K+1).$$

2.2 Steady-State Probability that 'i' Service Stations are Operating and 'j' Customers are in the System

Since we have assumed that the customers' interarrival times and the service times of two undistinguishable service stations are independently and exponentially distributed, the number of customers in the system has Markovian properties. From these Markovian properties, we can deduce that the mean transition rate into a given state is equal to the mean transition rate out of that state. Hence, we can write a set of steady-state probability equations which will lead to expressions for the probabilities of being in any of admissible states. Define $P(i,j)$ where $i=0,1,2.$ and $j=0,1,2,3,\cdots$ as the steady-state probability that 'i' service stations are operating and 'j' customers are in the system. Referring to the rate diagram for the M/M/2 system operating under the (0,K,N,M) policy appearing in Figure 2.1, the admissible states of the system can be partitioned into several disjoint subsets which lead to the following equations:

$$\lambda P(i,j) = \lambda P(i.j+1) \qquad (i=0,\ 0\leq j\leq N-2), \tag{2.1}$$

$$\lambda P(i,j) = \mu P(i+1,j+1) \qquad (i=0,\ j=0), \tag{2.2}$$

$$(\lambda+\mu)P(i,j) = \mu P(i,j+1) \qquad (i=1,\ j=1), \tag{2.3}$$

$$(\lambda+\mu)P(i,j) = \mu P(i,j+1) + \lambda P(i,j-1) \qquad (i=1,\ 2\leq j\leq M-2,\ j\neq K,N), \tag{2.4}$$

$$(\lambda+\mu)P(i,j) = \mu P(i,j+1) + \lambda P(i,j-1) + 2\mu P(i+1,j+1) \qquad (i=1,\ j=K), \tag{2.5}$$

$$(\lambda+\mu)P(i,j) = \mu P(i,j+1) + \lambda P(i,j-1) + \lambda P(i-1,j-1) \qquad (i=1,\ j=N), \tag{2.6}$$

$$(\lambda+\mu)P(i,j) = \lambda P(i,j-1) \qquad (i=1,\ j=M-1), \tag{2.7}$$

$$(\lambda+2\mu)P(i,j) = 2\mu P(i,j+1) \qquad (i=2,\ j=K+1), \tag{2.8}$$

$$(\lambda+2\mu)P(i,j) = 2\mu P(i,j+1) + \lambda P(i,j-1) \qquad (i=2,\ K+2\leq j,\ j\neq M), \tag{2.9}$$

$$(\lambda+2\mu)P(i,j) = 2\mu P(i,j+1) + \lambda P(i,j-1) + \lambda P(i-1,j-1) \qquad (i=2,j=M). \tag{2.10}$$

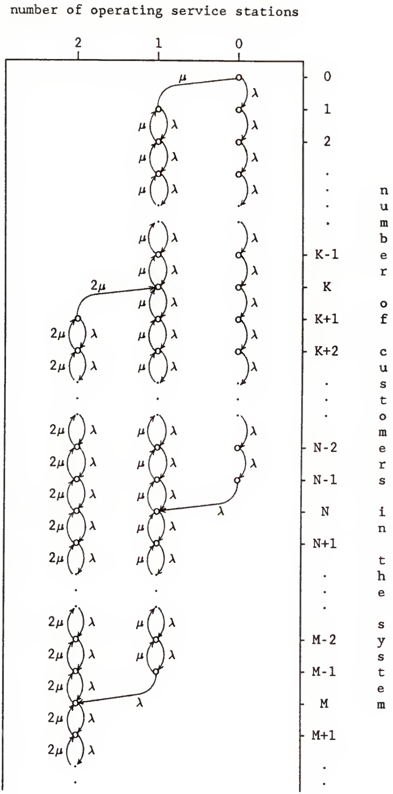

Figure 2.1 The rate diagram of the model with an infinite system capacity and an infinite population size

Define $\rho = \lambda/\mu$. P(i,j) for i=0,1,2. and j=1,2,3,$\cdots$ are obtained in terms of P(0,0) by solving equations (2.1) to (2.10) recursively. The summary of the results is given below.

For $\rho \neq 1$

$$
P(i,j) = \begin{cases}
P(0,0) & (i=0,\ 1 \le j \le N-1) \\[2mm]
\dfrac{\rho(1-\rho^{j})}{(1-\rho)} P(0,0) & (i=1,\ 1 \le j \le K) \\[4mm]
\left[\dfrac{\rho(1-\rho^{K})}{(1-\rho)} + \dfrac{\rho(\rho^{K}-\rho^{M-N})(1-\rho^{j-K})}{(1-\rho)(1-\rho^{M-K})} \right] P(0,0) & (i=1,\ K+1 \le j \le N) \\[4mm]
\dfrac{\rho(1-\rho^{N})(\rho^{j-N}-\rho^{M-N})}{(1-\rho)(1-\rho^{M-K})} P(0,0) & (i=1,\ N+1 \le j \le M-1) \\[4mm]
\dfrac{\rho^{M-N+1}(1-\rho^{N})[1-(\rho/2)^{j-K}]}{(2-\rho)(1-\rho^{M-K})} P(0,0) & (i=2,\ K+1 \le j \le M) \\[4mm]
\dfrac{\rho^{M-N+1}(1-\rho^{N})[1-(\rho/2)^{M-K}](\rho/2)^{j-M}}{(2-\rho)(1-\rho^{M-K})} P(0,0) & (i=2,\ j \ge M+1),
\end{cases}
\tag{2.11}
$$

and for $\rho = 1$

$$
P(i,j) = \begin{cases}
P(0,0) & (i=0,\ 1 \le j \le N-1) \\[2mm]
jP(0,0) & (i=1,\ 1 \le j \le K) \\[2mm]
\dfrac{NK + (M-K-N)j}{M-K} P(0,0) & (i=1,\ K+1 \le j \le N) \\[4mm]
\dfrac{N(M-j)}{M-K} P(0,0) & (i=1,\ N+1 \le j \le M-1) \\[4mm]
\dfrac{N[1-(1/2)^{j-K}]}{M-K} P(0,0) & (i=2,\ K+1 \le j \le M)
\end{cases}
\tag{2.12}
$$

$$P(i,j) = \left[\frac{N[1-(1/2)^{M-K}](1/2)^{j-M}}{M-K} P(0,0) \quad (i=2,\ j\geq M+1). \right. \tag{2.12}$$

Evaluation of P(0,0) can be done by the normalizing condition

$$\sum_{j=0}^{N-1} P(0,j) + \sum_{j=1}^{M-1} P(1,j) + \sum_{j=K+1}^{\infty} P(2,j) = 1.$$

In order for the left hand side of equation above to converge, ρ should be less than two. The normalizing condition can also be obtained from the probability generating function of the number of customers in the system. The evaluation of P(0,0) is deferred until the probability generating function of the number of customers in the system is derived in terms of P(0,0).

2.3 Probability Generating Function of the Number of Customers in the System

Define the probability generating function of the number of customers in the system, G(z) as

$$G(z) = \sum_{j=0}^{N-1} z^j P(0,j) + \sum_{j=1}^{M-1} z^j P(1,j) + \sum_{j=K+1}^{\infty} z^j P(2,j). \tag{2.13}$$

Using the steady-state probability equation (2.1), multiply by z^j and sum over all possible values of j. Then,

$$\sum_{j=0}^{N-1} z^j P(0,j) = \sum_{j=0}^{N-1} z^j P(0,0)$$

$$= \frac{1-z^N}{1-z} P(0,0). \tag{2.14}$$

Similarly, using equations (2.2) to (2.7),

$$\rho zP(0,0) + \sum_{j=2}^{M} z^j(1+\rho)P(1,j-1) =$$

$$+ \sum_{j=1}^{M-1} z^j P(1,j) + \sum_{j=3}^{M} \rho z^j P(1,N-2)$$

$$+ 2z^{K+1}P(2,K+1) + \rho z^{N+1}P(0,N-1).$$

Hence,

$$\sum_{j=1}^{M-1} z^j P(1,j) = \frac{\rho zP(0,0) - \rho z^{N+1}P(0,0)}{(1-\rho z)(1-z)}$$

$$+ \frac{\rho z^{M+1}P(1,M-1) - 2z^{K+1}P(2,K+1)}{(1-\rho z)(1-z)}. \tag{2.15}$$

Finally, for $\sum_{j=K+1}^{\infty} z^j P(2,j)$, using equations (2.8) to (2.10) yields

$$\sum_{j=K+1}^{\infty} z^j(1+\rho/2)P(2,j) = \sum_{j=K+1}^{\infty} z^j P(2,j+1)$$

$$+ \sum_{j=K+2}^{\infty} z^j(\rho/2)P(2,j-1) + z^M(\rho/2)P(1,M-1).$$

Thus,

$$\sum_{j=K+1}^{\infty} z^j P(2,j) = \frac{2z^{K+1}P(2,K+1) - \rho z^{M+1} P(1,M-1)}{(2-\rho z)(1-z)}. \tag{2.16}$$

Substituting (2.14), (2.15) and (2.16) into relation (2.13) and simplifying yield

$$G(z) = \frac{(1-z^N)}{(1-z)(1-\rho z)} P(0,0) + \frac{\rho z^{M+1}P(1,M-1) - 2z^{K+1}P(2,K+1)}{(1-z)(1-\rho z)(2-\rho z)}. \tag{2.17}$$

From the steady-state probabilities obtained in (2.11), P(1,M-1) and P(2,K+1) are given by

$$P(1,M-1) = \frac{\rho^{M-N}(1-\rho^{N})}{(1-\rho^{M-K})} P(0,0), \tag{2.18}$$

$$P(2,K+1) = \frac{\rho^{M-N+1}(1-\rho^{N})}{2(1-\rho^{M-K})} P(0,0). \tag{2.19}$$

Substituting the results for P(1,M-1) and P(2,K+1) in (2.18) and (2,19) into (2.17) and simplifying yield

$$G(z) = \frac{[(1-z^{N})(2-\rho z)(1-\rho^{M-K}) + \rho^{M-N+1}(1-\rho^{N})(z^{M+1}-z^{K+1})]P(0,0)}{(1-\rho^{M-K})(1-z)(1-\rho z)(2-\rho z)}. \tag{2.20}$$

The normalizing condition is obtained by using the property

$$G(z)\Big|_{z=1} = 1.$$

Apply L'Hospital's rule twice,

$$G(z)\Big|_{z=1} = \frac{-N(1-\rho^{M-K})(2-\rho) + \rho^{M-N+1}(1-\rho^{N})(M-K)}{-(1-\rho^{M-K})(1-\rho)(2-\rho)} P(0,0).$$

Hence,

$$P(0,0) = \begin{cases} \dfrac{(1-\rho^{M-K})(1-\rho)(2-\rho)}{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^{N})(M-K)} & \text{for } \rho \neq 1 \\[2ex] \dfrac{2}{N(M+K-N+4)}. & \text{for } \rho = 1. \end{cases} \tag{2.21}$$

By substituting the value above for P(0,0) in (2.21) into equation (2.20), it immediately follows that

$$G(z) = \frac{(1-\rho)(2-\rho)[(1-\rho^{M-K})(1-z^{N})(2-\rho z) - \rho^{M-N+1}(1-\rho^{N})(z^{M+1}-z^{K+1})]}{[N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^{N})(M-K)](1-z)(1-\rho z)(2-\rho z)}. \tag{2.22}$$

2.4 Expected Number of Customers in the System

Let J be the number of customers in the system. E[J] can be obtained explicitly from

$$E[J] = \left.\frac{dG(z)}{dz}\right|_{z=1}.$$

Hence, we obtain E[J] as

for $\rho \neq 1$

$$E[J] = \frac{M+K+1}{2} + \frac{\rho(3-2\rho)}{(1-\rho)(2-\rho)} - \frac{N(1-\rho^{M-K})\{(M+K-N)(2-\rho)+4\}}{2\{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(M-K)(1-\rho^{N})\}}, \tag{2.23}$$

for $\rho = 1$

$$E[J] = \frac{M+3}{3} + \frac{6(N-1)+(K+N)(K-N)+M(N-4)}{3(M+K-N+4)}. \tag{2.24}$$

2.5 Expected Number of Operating Service Stations

Under steady-state conditions, define the random variables O, J_q, W and W_q as

1) O ; the number of operating service stations,

2) J_q ; the number of customers who are waiting for service in the system (J_q does not include the customers who are being served.),

3) W : the customer's waiting time in the system (W does include service time.),

4) W_q ; the customer's waiting time for service (W_q does not include the service time.).

Then, the following relations hold:

$$E[J] = E[J_q] + E[O],$$

$$E[W] = E[W_q] + E[S_i].$$

From Little's formulas, we have

$$E[J] = \lambda E[W] \quad \text{and} \quad E[J_q] = \lambda E[W_q].$$

Hence, we obtain E[O] as

$$\begin{aligned} E[O] &= E[J] - E[J_q] \\ &= \lambda\{E[W] - E[W_q]\} \\ &= \lambda E[S_i] \\ &= \rho. \end{aligned} \tag{2.25}$$

Even though $E[J_q]$, $E[W]$ and $E[W_q]$ are not incorporated in the total expected cost function per unit time, they can be obtained explicitly from the above relations and the Little's formulas such as

$$E[J_q] = E[J] - E[O], \tag{2.26}$$

$$E[W] = \frac{1}{\lambda} E[J], \tag{2.27}$$

$$E[W_q] = \frac{1}{\lambda} E[J_q]. \tag{2.28}$$

2.6 Expected Length of the Idle Period, the Busy Period and the Busy Cycle

This removable queueing system alternates between durations when both service stations remain inoperative temporarily and durations when at least one service station is operating. To clarify the system states, the idle period, the busy period and the busy cycle are defined as follows.

1) Idle period denoted by I : the length of time when both service stations remain inoperative temporarily; this begins the instant when the last customer departs from the system and endures until the arrival of the next Nth customer.

2) Busy period denoted by B : the length of time when at least one service service station is operating; this begins the instant when the first service station starts operating at which time the number of waiting customers reaches N and both service stations remain inoperative temporarily, and lasts until there are no customers in the system. In other words, the busy period is the length of time between the end of the last idle period and the beginning of the next idle period.

3) Busy cycle denoted by C : the length of time between the beginning of last idle period and the beginning of the next idle period. In other word, the busy cycle is the sum of the idle period and the busy period such as

$$C = I + B, \tag{2.29}$$

or

$$E[C] = E[I] + E[B]. \tag{2.30}$$

After both service stations become inoperative temporarily, no service stations may operate until the number of customers who are waiting for service in the system reaches N. Hence, the idle period can be represented by the sum of the A_j's ($j=1,2\cdots,N$) which are the (j-1)th and the jth customers' interarrival times; thus,

$$I = A_1 + A_2 + A_3 + \cdots + A_{N-1} + A_N. \tag{2.31}$$

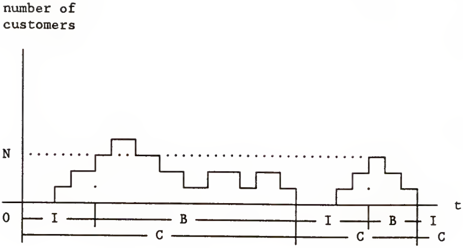

Figure 2.2 The sample function of the busy cycle

By the basic assumptions of system operation, the A_j's are identically, independently and exponentially distributed random variables with $E[A_j] = 1/\lambda$. Thus, the expected length of idle period, E[I] is obtained from

$$E[I] = \sum_{j=1}^{N} E[A_j] = N\, E[A_j] = N/\lambda. \tag{2.32}$$

Let P[O=0] be the probability that no service stations are operating. Then, P[O=0] can be computed from

$$P[O=0] = \sum_{j=0}^{N-1} P(0,j).$$

Hence,

$$P[O=0] = \begin{cases} \dfrac{N(1-\rho)(2-\rho)(1-\rho^{M-K})}{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^{N})(M-K)} & \text{for } \rho \neq 1 \\[2ex] \dfrac{2}{M+K-N+4} & \text{for } \rho = 1. \end{cases} \tag{2.33}$$

Since

$$P[O=0] = \frac{E[I]}{E[C]} = \frac{E[I]}{E[I] + E[B]},$$

then

$$E[B] = \frac{E[I] - E[I]\, P[O=0]}{P[O=0]}.$$

Therefore,

$$E[B] = \begin{cases} \dfrac{1}{\lambda}\left[\dfrac{N\rho}{(1-\rho)} - \dfrac{\rho^{M-N+1}(1-\rho^{N})(M-K)}{(1-\rho)(2-\rho)(1-\rho^{M-K})}\right] & \text{for } \rho \neq 1 \\[2ex] \dfrac{N(M+K-N+2)}{2\lambda} & \text{for } \rho = 1. \end{cases} \tag{2.34}$$

Substituting the explicit forms for E[I] in (2.32) and E[B] in (2.34) into relation (2.30) yields

$$E[C] = \begin{cases} \dfrac{1}{\lambda}\left[\dfrac{N}{(1-\rho)} - \dfrac{\rho^{M-N+1}(1-\rho^{N})(M-K)}{(1-\rho)(2-\rho)(1-\rho^{M-K})}\right] & \text{for } \rho \neq 1 \\[2ex] \dfrac{N(M+K-N+4)}{2\lambda} & \text{for } \rho = 1. \end{cases} \tag{2.35}$$

2.7 Expected Number of Activations and Removals of Service Stations per Unit Time

Define the random variables R_a and R_r as positive integers to be the number of activations and removals of service stations per unit time.

One of the service stations removed from the system should be activated immediately (i) whenever the number of customers in the system reaches N while both service stations remain inoperative temporarily, or (ii) whenever the number of customers in the system reaches M while one service stations are operating and the other remains inoperative. Since the assumed customers' mean arrival rate per is λ, $E[R_a]$ is obtained from

$$E[R_a] = \lambda P(0,N-1) + \lambda P(1,M-1). \tag{2.37}$$

Substituting P(0,N-1) and P(1,M-1) in (2.11) for $\rho \neq 1$ and in (2.12) for $\rho = 1$ with P(0,0) in (2.21) and simplifying yield

$$E[R_a] = \begin{cases} \dfrac{\lambda(1-\rho)(2-\rho)[(\rho^{M-N}-\rho^{M}) + (1-\rho^{M-K})]}{N(2-\rho)(1-\rho^{M-K}) - \rho^{M-N+1}(1-\rho^{N})(M-K)} & \text{for } \rho \neq 1 \\[2ex] \dfrac{2\lambda(M+N-K)}{N(M-K)(M+K-N+4)} & \text{for } \rho = 1. \end{cases} \tag{2.38}$$

Similarly, the service station just finishing service should be removed from the system immediately (i) whenever the number of customers in the system reaches 0 while one service station is operating and the other is removed from the system, or (ii) whenever the number of customers in the system reaches K while both service stations are operating simultaneously. However, mean service rate of each operating service station per unit time is μ. Thus, $E[R_r]$ can be obtained from

$$E[R_r] = \mu P(1,1) + 2\mu P(2,K+1). \tag{2.39}$$

Substituting the explicit values of P(1,1) and P(2,K+1) in (2.11) and (2.12) into relation (2.39) yields $E[R_r]$ explicitly which is the same as the results of $E[R_a]$ obtained in (2.38).

In the next chapter, we will analyze the busy period systematically to provide basic information to derive the probability mass function of the number of activations or removals of service stations during the busy period, the Laplace transform of the probability density function of the busy period and so on.

CHAPTER 3

ANALYSIS OF THE BUSY PERIOD

3.1 Introduction

In this chapter, we analyze the busy period systematically to provide basic information for (1) obtaining the probability mass function of the number of activations or removals of service station during the busy period, (2) deriving the Laplace transform of the probability density function of the busy period, (3) deriving the Laplace transforms of the probability density functions of the length of time when one service station is operating and of the length of time when both service stations are operating simultaneously and (4) developing alternative methodologies to obtain certain system characteristics that will be used to construct a total expected cost function in Chapter 7, without using the explicit form of $P(i,j)$.

The methodology used in this chapter is to partition the busy period into several random intervals depending on the numbers of customers and the number of operating service stations in the system. In other words, any busy period can be represented in terms of these random intervals. Derivation of the relation between the busy period and the random intervals is a prime objective of this chapter. We also analyze each random interval based on the fluctuation of the number of customers in the system. The properties of the random walk or the

classical gambler's ruin problem are used to characterize the fluctuation of the number of customers in the system.

3.2 Preliminary Work for Analysis of the Busy Period

To proceed towards a formal analysis of the busy period, define the random variables as

1) F_a and F_r ; both positive integers to be the number of activations and removals of service stations during the busy period,
2) D ; a nonnegative integer to be the number of times the level of customers in the system reaches M while one service station is operating during the busy period.

The main objective of this section is to show that

i) $F_a = F_r$,

ii) $F_a = D + 1$.

In any busy cycle, termination of the idle period and initiation of the busy period occur at the same time when the number of customers who are waiting for service reaches N while both service stations remain inoperative temporarily. This indicates that the busy period is initiated when one of the two service stations removed from the system starts operating and lasts until both service stations become inoperative temporarily again. Since the system operates under steady-state conditions, any service station which is operating during the busy period should be removed before the beginning of the next idle period. Thus, the number of activations of service stations is equal to the number of removals of service stations during the busy period. Hence, we have

$$F_a = F_r. \tag{3.1}$$

When one service station is operating and one service station is removed from the system, the level of customers may fluctuate between 1 and (M-1) without having the second service station activated, or after reaching M, the level fluctuates until it reaches K at which time only one service station is operating, after which time the level M is reached again and so on. Hence, there are no specific upper restrictions of fluctuations in the level of customers when both service stations are operating simultaneously. This indicates that the number of customers in the system can be allowed to increase indefinitely without changing the number of operating service stations. Therefore, the activations or the removals of the service stations depend on when the number of customers in the system reaches specified levels and how many service stations are operating. Thus, the following theorem might be stated based on the fluctuation of the number of customers in the system depending on the number of operating service stations.

Theorem 3.1

The number of activations or removals of the service stations during the busy period is equal to D + 1, i.e.,

$$F_a = D + 1 \qquad \text{for all } D. \tag{3.2}$$

Proof : Depending on the values of D, one of the following cases can occur during the busy period.

Case 1 : D = 0

Only the service station which is activated at the beginning of the busy period is operating during the entire busy period. The other available service station remains removed from the system during the entire busy

period. This case occurs only when the number of customers in the system is always less than M during the busy period. In this case, the number of activations or removals of the service stations during the busy period is 1, i.e.,

$$F_r = F_a = 1.$$

Case 2 : D = 1

One of the two service stations removed from the system starts operating at the beginning of the busy period. The remaining service station removed from the system also starts operating when the number of customers in the system reaches M during the busy period. Hence, the number of operating service stations becomes two. At a later time, when the number of customers in the system reaches K, the service station just finished service is removed from the system immediately. After this time, the number of customers in the system never reaches M again. Therefore, the busy period is terminated when the remaining service station is removed from the system. In this case, the number of activations or removals of service stations during the busy period is 2, i.e.,

$$F_r = F_a = 2.$$

Case 3 : $D \geq 2$

One of the two service stations removed from the system starts operating at the beginning of the busy period. The remaining service station removed from the system also starts operating when the number of customers in the system reaches M during the busy period. Hence, the number of operating service stations becomes two for the first time during the busy period. At a later time, when the number of customers in the system reaches K, the service station just finished service is

removed immediately. Thus, the number of operating service stations becomes one again. In other words, when the number of customers in the system reaches M from K without reaching 0 while one service station is operating, the number of operating service stations becomes two. Then, when the number of customers in the system reaches K from M without reaching 0, the number of operating service stations becomes one again. Thus, the number of operating service stations alternates between one and two (D-1) times. After that, the number of customers in the system never reaches M again. Hence, the busy period is terminated when the remaining operating service station is removed. In this case, the number of activations or removals of the service stations during the busy period is (D + 1), i.e.,

$$F_r = F_a = D + 1.$$

Hence, from the above three possible cases, the theorem follows immediately. Q.E.D.

From (3.2), we have

$$E[F_r] = E[F_a] = E[D] + 1. \tag{3.3}$$

Next, we will partition the busy period into the random intervals based on the number of customers and the number of operating service stations in the system.

3.3 Partition of the Busy Period

In order to represent mathematically the fluctuations of the number of customers in the system during the busy period, define the following random time intervals.

1) U_0 : the length of time from the beginning of the busy period to the first time the number of customers in the system reaches 0 without reaching M. Hence, U_0 is the first passage time from N to 0 without reaching M. During U_0, the level of customers in the system may fluctuate between 1 and (M-1) before it reaches 0 without reaching M given that one service station is operating.

2) U_1 : the length of time from the beginning of the busy period to the first time the number of customers in the system reaches M without reaching 0. Hence, U_1 is the first passage time from N to M without reaching 0. During U_1, the level of customers in the system may fluctuate between 1 and (M-1) before it reaches M without reaching 0 given that one service station is operating.

3) U_2 : the length of time from the instant when the number of customers in the system reaches K to the end of the busy period without reaching M. Hence, U_2 is the first passage time from K to 0 without reaching M. During U_2, the level of customers in the system may fluctuate between 1 and (M-1) before it reaches 0 without reaching M given that one service station is operating.

4) Y_d where $d=1,2,\cdots,D-1$ and $D \geq 1$: the dth length of time from the instant when the number of customers in the system reaches K while both service stations are operating to the first time the number of customers in the system reaches M without reaching 0. Hence, Y_d is the first passage time from K to M without reaching 0. During the Y_d's, the level of customers in the system may fluctuate between 1 and (M-1) before it reaches M without reaching 0 given that one service station is operating. Note that the Y_d's are identically and independently distributed and independent of D and F_a.

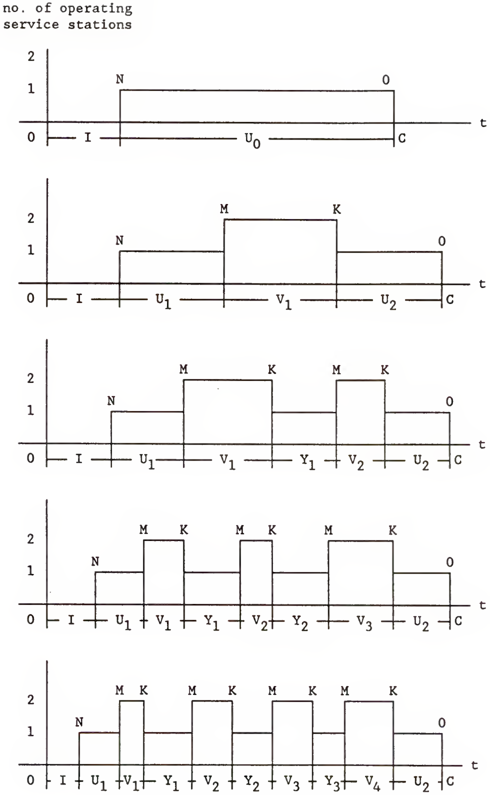

Figure 3.1 The sample functions of the fluctuations of the number of customers in the system

5) V_d where $d=1,2,\cdots,D$ and $D \geq 1$: the dth length of time from the instant when the number of customers in the system reaches M while one service station is operating to the first time the number of customers in the system reaches K. Hence, V_d is the first passage time from M to K. During the V_d's, the level of customers in the system may fluctuate between (K+1) and an indefinite level which is greater than or equal to M before it reaches K given that both service stations are operating simultaneously. Note that the V_d's are identically and independently distributed and independent of D and F_a.

We note in particular that the length of any busy period may be represented in terms of the random variables U_0, U_1, U_2, V_d and Y_d for all d.

Theorem 3.2

Any busy period, B can be expressed in the following form:

$$B = \begin{cases} U_0 & \text{for } D = 0 \\ U_1 + V_1 + U_2 & \text{for } D = 1 \\ U_1 + \sum_{d=1}^{D-1} Y_d + \sum_{d=1}^{D} V_d + U_2 & \text{for } D \geq 2. \end{cases} \tag{3.4}$$

Proof : Theorem 3.2 follows the interpretation of the sample functions of the fluctuations of the number of customers in the system depending on the number of operating service stations appearing in Figure 3.1.

Theorem 3.2 is the basic information for (i) obtaining the probability mass function of F_a, (ii) deriving the Laplace transform of the probability density function of the busy period, B and (iii)

developing alternative methodologies to obtain certain system characteristics without using the explicit form of $P(i,j)$.

However, during U_0, U_1, U_2 and Y_d for all d, exactly one service station is operating. During V_d for all d, both service stations are operating simultaneously. Hence, any busy period can be represented by the sum of the length of time when one service station is operating and the length of time when both service stations are operating simultaneously. Let B_1 and B_2 be the random variables which represent the length of time when one service station is operating and the length of time when both service stations are operating simultaneously during the busy period, respectively. Then, we have

$$B = B_1 + B_2. \tag{3.5}$$

Hence, from relation (3.4) and (3.5), B_1 and B_2 can also be represented in terms of U_0, U_1, U_2, Y_d and V_d for all d.

<u>Lemma 3.1</u>

i) The length of time when exactly one service station is operating during the busy period, B_1 can be expressed by

$$B_1 = \begin{cases} U_0 & \text{for } D = 0 \\ U_1 + U_2 & \text{for } D = 1 \\ U_1 + \sum_{d=1}^{D-1} Y_d + U_2 & \text{for } D \geq 2. \end{cases} \tag{3.6}$$

ii) The length of time when both service stations are operating simultaneously during the busy period can be expressed by

$$B_2 = \begin{cases} V & \text{for } D = 0 \\ \sum_{d=1}^{D} V_d & \text{for } D \geq 1 \end{cases} \tag{3.7}$$

where V is the random variable with

$$P[V=0] = 1.$$

Next, we will characterize U_0, U_1, U_2, Y_1 and V_1 based on the fluctuation of the number of customers in the system in order to characterize the busy period.

3.4 Characterizations of U_0, U_1, U_2, Y_1 and V_1

During a busy period, the number of customers in the system is increased by 1 or decreased by 1 instantly due to a transition which could be a new customer's arrival or a customer's departure from the system. However, after the number of customers in the system reaches x during the busy period, the system state x sojourns for a random length of time until it becomes (x+1) or (x-1). Thus, the length of time from the instant the number of customers in the system reaches x to the first time the number of customers in the system reaches y can be characterized by both (i) the random length of sojourn time between two successive transitions and (ii) the random number of transitions until the number of customers in the system reaches y for the first time. To do so, define the following random variables:

1) U ; the length of time from the instant the number of customers in the system reaches x to the first time the number of customers in the system reaches y during a busy period,

2) L ; the number of transitions during U where L is a nonnegative integer,

3) T_n ; the length of sojourn time between the (n-1)th and the nth transitions during U where $n=1,2,\cdots,L$,

4) J(n) ; the number of customers in the system just after the nth transition during U where $n=1,2,\cdots,L$.

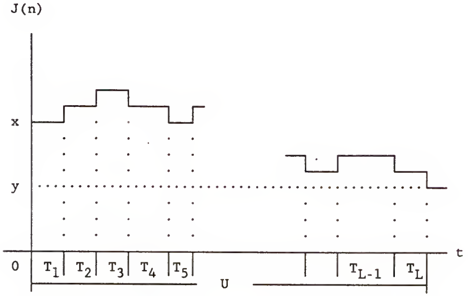

Figure 3.2 The fluctuation of the number of customers in the system during U

Hence, U can be expressed in terms of T_n ($n=1,2,\cdots,L$) such as

$$U = T_1 + T_2 + T_3 + \cdots + T_{L-1} + T_L. \tag{3.8}$$

In a Markov process, every time epoch at which a transition takes place is a regeneration point which implies that development of future process is independent of the history of process (see Prabhu[20]). Hence, the T_n's are identically distributed and independent of L. Therefore, for any given system states x and y, U can be characterized from relation (3.8) if the random variables T_n and L are evaluated.

3.4.1 Evaluation of the Length of Time Between Two Successive Transitions during U

Using the following theorem, the random variable T_n for all n can be evaluated.

Theorem 3.3

1) If one service station is operating, then T_n is exponentially distributed with parameter $\mu(1+\rho)$.

2) If both service stations are operating simultaneously, then T_n is exponentially distributed with parameter $\mu(2+\rho)$.

Proof : From the basic assumptions of system operation in Chapter 1, the random variables A_1, S_1 and S_2 are exponentially and independently distributed with parameter λ, μ and μ, respectively.

1) Since $T_1 = \min\{A_1, S_1\}$, we have

$$
\begin{aligned}
P[T_1>t] &= P[\min(A_1,S_1) > t] \\
&= P[A_1>t\ ,\ S_1>t] \\
&= P[A_1>t]\ P[S_1>t] \\
&= e^{-\lambda t}\ e^{-\mu t} \\
&= e^{-(\lambda+\mu)t}.
\end{aligned}
$$

2) Similarly, using $T_1 = \min\{A_1, S_1, S_2\}$ yields

$$
\begin{aligned}
P[T_1>t] &= P[\min(A_1,S_1,S_2) > t] \\
&= P[A_1>t\ ,\ S_1>t\ ,\ S_2>t] \\
&= P[A_1>t]\ P[S_1>t]\ P[S_2>t] \\
&= e^{-\lambda t}\ e^{-\mu t}\ e^{-\mu t} \\
&= e^{-(\lambda+2\mu)t}.
\end{aligned}
$$

Thus, the theorem follows immediately. Q.E.D.

3.4.2 Evaluation of the Number of Transitions during U

In the stochastic process $\{J(n),\ n=1,2,\cdots,L\}$, the fluctuations of the number of customers in the system is a random walk. Hence, L, the number of transitions during U is dictated by the one step transition probabilities given the initial and the final levels of the number of customers in the systems. Let the one step transition probability matrix of J(n) during U be H and also let

$$p = P[J(n+1)=j+1|J(n)=j]$$

and

$$q = P[J(n+1)=j-1|J(n)=j],$$

where state j is not the final level of the number of customers in the system, y during U.

Since

$$p + q = 1, \tag{3.9}$$

we have

$$P[J(n+1)=j|J(n)=j] = 0.$$

If state j is the final level of the number of customers in the system, y which forms an absorbing barrier, then

$$p = P[J(n+1)=y+1|J(n)=y] = 0,$$

$$q = P[J(n+1)=y-1|J(n)=y] = 0,$$

$$P[J(n+1)=y|J(n)=y] = 1.$$

In general, H exhibits a form with an absorbing barrier at y if we assume without loss of generality x>y, appearing in Figure 3.3.

$$
\begin{array}{c|cccccccccc|}
J(n) & y & y+1 & y+2 & \cdots & x-2 & x-1 & x & x+1 & x+2 & \cdots \\
\hline
y & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
y+1 & q & 0 & p & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
y+2 & 0 & q & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \\
x-2 & 0 & 0 & 0 & \cdots & 0 & p & 0 & 0 & 0 & \cdots \\
x-1 & 0 & 0 & 0 & \cdots & q & 0 & p & 0 & 0 & \cdots \\
x & 0 & 0 & 0 & \cdots & 0 & q & 0 & p & 0 & \cdots \\
x+1 & 0 & 0 & 0 & \cdots & 0 & 0 & q & 0 & p & \cdots \\
x+2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & q & 0 & \cdots \\
\vdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots
\end{array}
$$

Figure 3.3 H, the one step transition probability matrix of the number of customers in the system during U

The values of p and q are obtained using the following theorem.

Theorem 3.4

1) If one service station is operating, then

$$p = \frac{\rho}{(1+\rho)} \quad \text{and} \quad q = \frac{1}{(1+\rho)}. \tag{3.10}$$

2) If both service stations are operating simultaneously, then

$$p = \frac{\rho}{(2+\rho)} \quad \text{and} \quad q = \frac{2}{(2+\rho)}. \tag{3.11}$$

Proof : 1) By Cinlar [7] and Taylor and Karlin[24], or

$$p = P[A_1<S_1] = \iint_{0<x<y<\infty} \lambda e^{-\lambda x} \mu e^{-\mu y}\, dxdy$$

$$= \frac{\lambda}{(\lambda+\mu)}$$

$$= \frac{\rho}{(1+\rho)}.$$

2) Similiarly,

$$p = P[A_1<S_1<S_2] + P[A_1<S_2<S_1]$$

$$= \iiint_{0<x<y<z<\infty} \lambda e^{-\lambda x} \mu e^{-\mu y} \mu e^{-\mu z}\, dxdydz + \iiint_{0<x<z<y<\infty} \lambda e^{-\lambda x} \mu e^{-\mu z} \mu e^{-\mu y}\, dxdzdy$$

$$= \frac{\lambda}{(\lambda+2\mu)}$$

$$= \frac{\rho}{2+\rho}.$$

From the above p's and relation (3.9), the theorem follows immediately.

Q.E.D.

Thus, the random variable L can be evaluated by the one step transition probability matrix of J(n), H appearing in Figure 3.3 with the values of p and q shown in (3.10) and (3.11).

Next, we will characterize U_0, U_1, U_2, Y_1 and V_1 based on the results of evaluations of T_n and L.

3.4.3 Characterizations of U_0, U_1, U_2 and Y_1

Let the nonnegative random variables L_0, L_1, L_2 and L_3 be the number of transitions during U_0, U_1, U_2 and Y_d for all d, respectively. Then, similar to U, we have

$$U_0 = T_1 + T_2 + T_3 + \cdots + T_{L_0}, \tag{3.12}$$

$$U_1 = T_1 + T_2 + T_3 + \cdots + T_{L_1}, \tag{3.13}$$

$$U_2 = T_1 + T_2 + T_3 + \cdots + T_{L_2}, \tag{3.14}$$

$$Y_1 = T_1 + T_2 + T_3 + \cdots + T_{L_3}. \tag{3.15}$$

During U_0, U_1, U_2 and Y_1, one service station is operating. Hence, the T_n's for all n are identically, independently and exponentially distributed with parameter $\mu(1+\rho)$. After the number of customers in the system reaches N when both service stations remain inoperative temporarily, the number of customers in the system might fluctuate between 1 and M-1 before it reaches 0 or M for the first time. Hence, one of the two possible cases will occur whether the number of customers in the system will reach 0 for the first time without reaching M which leads to U_0, or M for the first time without reaching 0 occurs which leads to U_1 after the number of customers in the system reaches N at the beginning of the busy period. The initial levels of the number of customers in the system in U_0 and U_1 are the same, namely N and the

final levels in U_0 and U_1 are 0 and M, respectively. Thus, L_0 and L_1, the number of transitions during U_0 and U_1 are dictated by the one step transition probability matrix of the number of customers in the system as shown in Figure 3.4.

Similarly, the one step transition probability matrix of J(n) during U_2 and Y_1 which governs L_2 and L_3, the number of transitions during U_2 and Y_1 can be identified. After the number of customers in the system reaches K for the first time while both service stations are operating simultaneously, the number of customers in the system will reach 0 for the first time without reaching M which leads to U_2, or M for the first time without reaching 0 which leads to Y_1. The initial levels of the number of customers in the system in U_2 and Y_1 are the same, namely K. But the final levels in U_2 and Y_1 are 0 and M, respectively. Thus, L_2 and L_3, the number of transitions during U_2 and Y_1 are also dictated by the same one step transition probability matrix of J(n) during U_0 and U_1 which governs L_0 and L_1 in Figure 3.4.

Comparing U_0 and U_2, both cases have the same final level of the number of customers in the system, namely 0 which is an absorbing barrier. L_0 and L_2, the number of transitions during U_2 and Y_1 are dictated by the same one step transition probability matrix of J(n) appearing in Figure 3.4. The only difference is the initial levels of the number of customers in the system, i.e., N in U_0 and K in U_2. This relation between U_0 and U_2 could be helpful in finding certain characteristics related to U_0 and U_2 from each other. Note that the same relation between U_1 and Y_1 also holds.

$$
\begin{array}{c|ccccccccccc|}
J(n) & 0 & 1 & 2 & \cdots & N-1 & N & N+1 & \cdots & M-2 & M-1 & M \\
0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
1 & q & 0 & p & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
2 & 0 & q & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\
N-1 & 0 & 0 & 0 & \cdots & 0 & p & 0 & \cdots & 0 & 0 & 0 \\
N & 0 & 0 & 0 & \cdots & q & 0 & p & \cdots & 0 & 0 & 0 \\
N+1 & 0 & 0 & 0 & \cdots & 0 & q & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\
M-2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & p & 0 \\
M-1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & q & 0 & p \\
M & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1
\end{array}
$$

Figure 3.4 The one step transition probability matrix of the number of customers in the system during U_0 and U_1 where the values of p and q are shown in (3.10)

3.4.4 Characterization of V_1

Define the nonnegative integer random variable L_4 as the number of transitions during V_1. Then, V_1 can also be represented in terms of T_n and L_4 as

$$V_1 = T_1 + T_2 + T_3 + \cdots + T_{L_4}. \tag{3.16}$$

During V_1, both service stations are operating simultaneously. Hence, the T_n's in relation (3.16) are distributed identically, independently and exponentially with paramenter $\mu(2+\rho)$. After the number of customers in the system reaches M while one service station is operating, the number of customers in the system might fluctuate between K+1 and an indefinite level which is greater than or equal to M. Then, eventually the number of customers in the systèm reaches K for the first time after a finite number of transitions because the system is operating under steady-state conditions. Thus, the one step transition probability matrix of J(n) during V_1 which governs L_4 is given by Figure 3.5.

Based on the results of characterizing U_0, U_1, U_2, Y_1 and V_1, it is possible to use relation (3.4) to proceed towards the objectives of this research.

In the next chapter, we will derive the probability mass function of the number of activations or removals of service stations during the busy using the results of this chapter.

$$
\begin{array}{c|ccccccccc|}
J(n) & K & K+1 & K+2 & \cdots & M-1 & M & M+1 & M+2 & \cdots \\
K & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots \\
K+1 & q & 0 & p & \cdots & 0 & 0 & 0 & 0 & \cdots \\
K+2 & 0 & q & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots \\
M-1 & 0 & 0 & 0 & \cdots & 0 & p & 0 & 0 & \cdots \\
M & 0 & 0 & 0 & \cdots & q & 0 & p & 0 & \cdots \\
M+1 & 0 & 0 & 0 & \cdots & 0 & q & 0 & p & \cdots \\
M+2 & 0 & 0 & 0 & \cdots & 0 & 0 & q & 0 & \cdots \\
\cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots \\
\cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots
\end{array}
$$

Figure 3.5 The one step transition probability matrix of the number of customers in the system during V_1 where the values of p and q are shown in (3.11)

CHAPTER 4

PROBABILITY MASS FUNCTION OF THE NUMBER OF ACTIVATIONS OR REMOVALS OF SERVICE STATIONS DURING THE BUSY PERIOD

4.1 Introduction

The main objective of this chapter is to derive the probability mass function of the number of activations F_a or removals F_r of service stations during the busy period. The explicit form of the probability mass function of F_a or F_r provides opportunities (1) to obtain the Laplace transforms of the probability density functions of the busy period, of the busy cycle, of the length of time when one service station is operating and of the length of time when both service stations are operating simultaneously during a busy period and (2) to develop alternative methodologies i) to find some of the system characteristics derived in Chapter 2 without using P(i,j) and ii) important system characteristics which may not be obtainable from P(i,j).

In order to find the probability mass function of F_a or F_r explicitly, we use the results of Chapter 3 and the properties of the classical gambler's ruin problem.

4.2 Derivation of the Probability Mass Function of Activation or Removals of Service Stations during the Busy Period

In Chapter 3, we showed that

$$F_a = D + 1 \quad \text{and} \quad F_r = D + 1. \tag{4.1}$$

Hence, relation (3.4) can be rewritten as

$$B = \begin{cases} U_0 & \text{for } F_a = 1 \\ U_1 + V_1 + U_2 & \text{for } F_a = 2 \\ U_1 + \sum_{r=1}^{F_a-1} V_r + \sum_{r=1}^{F_a-2} Y_r + U_2 & \text{for } F_a \geq 3. \end{cases} \tag{4.2}$$

From relation (4.2), it is evident that there is a one-to-one correspondence between the value of F_a and the distribution of the busy period, B. In other words, for any given value of F_a, the distribution of B is uniquely determined in terms of U_0, U_1, U_2, Y_r and V_r for all r. This implies that the probability that the number of activations or removals of service stations during the busy period is r is equal to the probability that the uniquely determined distribution of the busy period given that F_a is r will occur. Hence, we define the following probabilities.

1) $P[F_a = r]$, $r=1,2,\cdots$; the probability that the number of activations of service stations during the busy period is r, i.e., the probability mass function of F_a or F_r.
2) $P(U_0)$; the probability that the number of customers in the system will reach 0 for the first time without reaching M after the number of customers in the system reaches N at the beginning of the busy period, or the probability that the number of customers in the system is always less than M during the entire busy period.

Equivalently, this is the probability that U_0 will occur after the number of customers in the system reaches N at the beginning of the busy period.

3) $P(U_1)$; the probability that the number of customers in the system will reach M for the first time without reaching 0 after the number of customers in the system reaches N at the beginning of the busy period. Equivalently, this is the probability that U_1 will occur after the number of customers in the system reaches N at the beginning of the busy period.

4) $P(U_2)$; the probability that the number of customers in the system will reach 0 for the first time without reaching M after the number of customers in the system reaches K when both service stations are operating simultaneously during the busy period. Equivalently, this is the probability that U_2 will occur after the number of customers in the system reaches K when both service station is operating simultaneously during the busy period.

5) $P(Y_1)$; the probability that the number of customers in the system will reach M for the first time without reaching 0 after the number of customers in the system reaches K when both service stations are operating during the busy period. Equivalently, this is the probability that Y_r for all r will occur after the number of customers in the system reaches K when both service station is operating simultaneously during the busy period.

6) $P(V_1)$; the probability that the number of customers in the system will reach eventually K for the first time after the number of customers in the system reaches M when one service station is operating during the busy period.

Hence, it is easy to deduce the following relations from the definitions of U_0, U_1, U_2, Y_d and V_d for all d under steady-state conditions:

$$P(U_0) + P(U_1) = 1, \tag{4.3}$$

$$P(U_2) + P(Y_1) = 1, \tag{4.4}$$

$$P(V_1) = 1. \tag{4.5}$$

Furthermore, according to relation (4.2), the following theorem may be stated based on the one-to-one correspondence between the value of F_a and the uniquely determined distribution of the busy period given the value of F_a.

Theorem 4.1

The probability mass function of F_a or F_r, $P[F_a=r]$ for $r=1,2,3,\cdots$ can be obtained from the following relation:

$$P[F_a=r] = \begin{cases} P(U_0) & \text{for } r = 1 \\ P(U_1)P(U_2)[P(V_1)]^{r-1}[P(Y_1)]^{r-2} & \text{for } r \geq 2, \end{cases} \tag{4.6}$$

or using relation (4.4), relation (4.6) is equivalent to

$$P[F_a=r] = \begin{cases} P(U_0) & \text{for } r = 1 \\ P(U_1)P(U_2)\ [P(Y_1)]^{r-2} & \text{for } r \geq 2. \end{cases} \tag{4.7}$$

Theorem 4.1 provides the fundamental relation to obtain $P[F_a=r]$ for all r; this depends only on $P(U_0)$, $P(U_1)$, $P(U_2)$ and $P(Y_1)$.

Our next objective is to derive explicit forms for the above four probabilities in terms of λ, μ, K, N and M in order to obtain $P[F_a=r]$ for all r explicitly. To do so, we use the properties of the gambler's

ruin problem because of the similarity between the fluctuations of the number of customers in the system in a Markovian death & birth process and the duration of a game in the gambler's ruin problem.

4.2.1 Evaluations of $P(U_0)$ and $P(U_1)$

Once the busy period is initiated when the number of customers in the system reaches N while both service stations remain inoperative temporarily, the number of customers in the system will reach 0 for the first time without reaching M with probability $P(U_0)$ or, M for the first time without reaching 0 with probability $P(U_1)$. However, the fluctuation of the number of customers in the system during U_0 and U_1 is dictated by the one step transition probability matrix of the number of customers in the system as exhibited in Figure 3.4. During U_0 and U_1, the number of operating service stations is one. The values of p and q, the elements of the one step transition probability matrix of J(n) are given by (3.10) such as

$$p = \frac{\rho}{1+\rho} \qquad \text{and} \qquad q = \frac{1}{1+\rho}. \tag{4.8}$$

In order to evaluate $P(U_0)$, let

$$\Phi_0(N) = P(U_0). \tag{4.9}$$

Then, $\Phi_0(N)$ satisfies the following second order difference equations:

$$\Phi_0(N) = p\,\Phi_0(N+1) + q\,\Phi_0(N-1) \qquad 1 \le N \le M-1 \tag{4.10}$$

with boundary conditions

$$\Phi_0(0) = 1 \tag{4.11}$$

$$\Phi_0(M) = 0. \tag{4.12}$$

The solution of the difference equation above

$$\Phi_0(N) = \begin{cases} \dfrac{(q/p)^M - (q/p)^N}{(q/p)^M - 1} & \text{for } p \neq q \\ \dfrac{M-N}{M} & \text{for } p = q. \end{cases} \tag{4.13}$$

Substituting the values of p and q in (3.10) or (4.8) and simplifying yield

$$P(U_0) = \Phi_0(N) = \begin{cases} \dfrac{1-\rho^{M-N}}{1-\rho^M} & \text{for } \rho \neq 1 \\ \dfrac{M-N}{M} & \text{for } \rho = 1. \end{cases} \tag{4.14}$$

Note that $p \neq q$ and $p = q$ are equivalent to $\rho \neq 1$ and $\rho = 1$.

$P(U_1)$ can be obtained from either solving a difference equation similar to (4.10) or using relation (4.3) and the results of $P(U_0)$ in (4.14). Hence, we obtain

$$P(U_1) = \begin{cases} \dfrac{\rho^{M-N} - \rho^M}{1-\rho^M} & \text{for } \rho \neq 1 \\ \dfrac{N}{M} & \text{for } \rho = 1. \end{cases} \tag{4.15}$$

4.2.2 Evaluations of $P(U_2)$ and $P(Y_1)$

Suppose the number of customers in the system reaches K when both service stations are operating simultaneously. Then, the number of customers in the system will reach 0 for the first time without reaching

M with probability $P(U_2)$, or M for the first time without reaching 0 with probability $P(Y_1)$. We may construct and solve in a similar way a second order difference equation to obtain $P(U_2)$ and $P(Y_1)$. However, we can deduce $P(U_2)$ and $P(Y_1)$ directly from the results of $P(U_0)$ and $P(U_1)$ in (4.14) and (4.15). One service station is operating during U_0 and U_2 and both U_0 and and U_2 are first passage times to reach 0 without reaching M. This leads to the same values of p and q as in (4.8), which are the elements of the one step transition probabilities of J(n). The only difference between U_0 and U_2 is the initial levels of the number of customers in the system, which are N in U_0 and K in U_2. Hence, $P(U_2)$ are obtained directly from the results of $P(U_0)$ by replacing N by K. Similar relation also holds between U_1 and Y_1. Thus, using the results of $P(U_0)$ and $P(U_1)$ in (4.14) and (4.15), we replace N by K to obtain

$$P(U_2) = \begin{cases} \dfrac{1-\rho^{M-K}}{1-\rho^M} & \text{for } \rho \neq 1 \\[2ex] \dfrac{M-K}{M} & \text{for } \rho = 1, \end{cases} \tag{4.16}$$

$$P(Y_1) = \begin{cases} \dfrac{\rho^{M-K}-\rho^M}{1-\rho^M} & \text{for } \rho \neq 1 \\[2ex] \dfrac{K}{M} & \text{for } \rho = 1. \end{cases} \tag{4.17}$$

Note that because of the similarities between the Markovian birth and death process and the classical gambler's ruin problem, $P(U_0)$, $P(U_2)$ $P(U_2)$ and $P(Y_1)$ can also be deduced from the gambler's ultimate ruin probabilities (see Feller[10], pp 344).

Since all four probabilities $P(U_0)$, $P(U_1)$, $P(U_2)$ and $P(Y_1)$ are derived, the probability mass function of the number of activations or removals during a busy period, $P[F_a=r]$ for $r=1,2,3,\cdots$ can be obtained using relation (4.7) and the explicit values of $P(U_0)$, $P(U_1)$, $P(U_2)$ and $P(Y_1)$.

4.2.3 Derivation of the Probability Mass Function of F_a

Substituting the explicit results of $P(U_0)$, $P(U_1)$, $P(U_2)$ and $P(Y_1)$ in (4.15) to (4.18) into relation (4.7) yields the probability mass function of F_a, $P[F_a=r]$ for all r as

for $\rho \neq 1$

$$P[F_a=r] = \begin{cases} \dfrac{1-\rho^{M-N}}{1-\rho^M} & \text{for } r = 1 \\[2ex] \dfrac{(1-\rho^{M-K})(\rho^{M-N}-\rho^M)}{(1-\rho^M)^2} \left[\dfrac{\rho^{M-K}-\rho^M}{1-\rho^M}\right]^{r-2} & \text{for } r \geq 2, \end{cases} \tag{4.18}$$

for $\rho = 1$

$$P[F_a=r] = \begin{cases} \dfrac{M-N}{M} & \text{for } r = 1 \\[2ex] \dfrac{N(M-K)}{M^2} \left[\dfrac{K}{M}\right]^{r-2} & \text{for } r \geq 2. \end{cases} \tag{4.19}$$

From relation (3.2) and $P[F_a=r]$ in (4.18) and (4.19), the probability mass function of D, denoted by $P[D = d]$ for $d= 0,1,2,\cdots$ is also obtained as

for $\rho \neq 1$

$$P[D=d] = \begin{cases} \dfrac{1-\rho^{M-N}}{1-\rho^{M}} & \text{for } d = 0 \\[2ex] \dfrac{(1-\rho^{M-K})(\rho^{M-N}-\rho^{M})}{(1-\rho^{M})^2} \left[\dfrac{\rho^{M-K}-\rho^{M}}{1-\rho^{M}}\right]^{d-1} & \text{for } d \geq 1, \end{cases} \tag{4.20}$$

for $\rho = 1$

$$P[D=d] = \begin{cases} \dfrac{M-N}{M} & \text{for } d = 0 \\[2ex] \dfrac{N(M-K)}{M^2} \left[\dfrac{K}{M}\right]^{d-1} & \text{for } d \geq 1. \end{cases} \tag{4.21}$$

It is easy to verify that

$$\sum_{r=1}^{\infty} P[F_a=r] = \sum_{d=0}^{\infty} P[D=d] = 1. \tag{4.22}$$

Hence, $P[F_a=r]$ and $P[D=d]$ in (4.18) to (4.21) are proper probability mass functions of F_a and D.

In the next chapter, we will derive the Laplace transforms of the probability mass function of the busy period using the result of the probability mass function of the number of activations or removals of service stations during the busy period.

CHAPTER 5

LAPLACE TRANSFORM OF THE PROBABILITY DENSITY FUNCTION OF THE BUSY PERIOD

5.1 Introduction

According to the results characterizing the random variables U_0, U_1, U_2, Y_1 and V_1 obtained in Chapter 3, each random variable depends on the number of transitions and the length of time between two successive transitions. The length of times between two successive transitions are identically, independently and exponentially distributed and independent of the number of transitions. Thus, it is possible to obtain the probability density functions of U_0, U_1, U_2, Y_1 and V_1 if the probability mass functions of the the number of transitions during U_0, U_1, U_2, Y_1 and V_1, respectively, are derived.

In this chapter, we first derive the probability generating functions of the number of transitions during U_0, U_1, U_2, Y_1 and V_1. Then, using these results, we proceed towards finding the Laplace transforms of the probability density functions of U_0, U_1, U_2, Y_1 and V_1 which are the basic components of the length of times B, C, B_1 and B_2 related to the busy period. Finally, we derive the Laplace transforms of the probability density functions of B, C, B_1 and B_2. The results of this chapter will be used in Chapter 6 to find certain system characteristics without using the joint probability mass function P(i,j).

5.2 Probability Generating Functions of the Number of Transitions during U_0, U_1, U_2, Y_1 and V_1

In order to find the probability generating functions of the number of transitions during U_0, U_1, U_2, Y_1 and V_1, we also use the properties of the classical gambler's ruin problem. Hence, let $P[L_0 = m]$ for $m=0,1,2,\cdots$ be the probability that the number of transitions during U_0 is m. Also let $G_0(N,z)$, $G_1(N,z)$, $G_0(K,z)$, $G_1(K,z)$ and $G_2(M,z)$ be the probability generating functions of L_0, L_1, L_2, L_3 and L_4, the number of transitions during U_0, U_1, U_2, Y_1 and V_1, respectively. For example,

$$G_0(N,z) = \sum_{m=0}^{\infty} z^m P[L_0 = m]. \tag{5.1}$$

In order to find the probability generating functions of the L_n's, we extensively use the probability generating function technique for the duration of a game, developed by Feller[10].

5.2.1 $G_0(N,z)$ and $G_1(N,z)$

Suppose that the number of customers in the system reaches N when both service stations remain inoperative temporarily. Then, the number of customers in the system fluctuates between 1 and M-1 before it will reach 0 for the first time without reaching M with probability $P(U_0)$, or M for the first time without reaching 0 with probability $P(U_1)$. The explicit values of $P(U_0)$ and $P(U_1)$ were already computed in (4.14) and (4.15) and given by

$$P(U_0) = \frac{1-\rho^{M-N}}{1-\rho^M}, \tag{5.2}$$

$$P(U_1) = \frac{\rho^{M-N}-\rho^M}{1-\rho^M}. \tag{5.3}$$

Note that

$$P(U_0) + P(U_1) = 1. \tag{5.4}$$

In order to compute $G_0(N,z)$, let $P_0(N,m)$ for $m=0,1,2,\cdots$ be the unconditional probability that the number of transitions is m from the epoch when the number of customers in the system reaches N at the beginning of the busy period to the epoch when this number reaches 0 without reaching M. Then, we have

$$\sum_{m=0}^{\infty} P_0(N,m) = P(U_0), \tag{5.5}$$

$$P[L_0 = m] = \frac{P_0(N,m)}{P(U_0)} \qquad \text{for all } m. \tag{5.6}$$

The fluctuation of the number of customers in the system is a random walk. Hence, in order to find $G_0(N,z)$, we use the methodology developed by Feller[10] for finding the probability generating function of the duration of a game in a classical gambler's ruin problem.

After the first transition at the beginning of a busy period, the number of customers in the system is N+1 with probability p, or N-1 with probability q. Thus, $P_0(N,m)$ satisfies the following second order difference equation:

$$P_0(N,m+1) = p\, P_0(N+1,m) + q\, P_0(N-1,m) \qquad \text{for } m \geq 0 \text{ and } 0 < N < M \tag{5.7}$$

with boundary conditions

$$P_0(0,m) = P_0(M,m) = 0 \qquad \text{for } m \geq 1 \tag{5.8}$$

$$P_0(N,0) = 0 \qquad \text{for } 0 < N \leq M \tag{5.9}$$

$$P_0(0,0) = 1 \tag{5.10}$$

where the values of p and q are shown in (3.10).

Note that the fluctuation of the number of customers in the system when one service station is operating during the busy period is dictated by the one step transition probability matrix of the number of customers in the system appearing in Figure 3.4.

Define

$$g_0(N,z) \;=\; \sum_{m=0}^{\infty} z^m \, P_0(N,m). \tag{5.11}$$

Multiplying (5.7) by z^{m+1} and adding over all possible values of m, we have

$$g_0(N,z) = pz\, g_0(N+1,z) + qz\, g_0(N-1,z) \qquad \text{for } 0 < N < M \tag{5.12}$$

with boundary conditions

$$g_0(0,z) \;=\; 1 \tag{5.13}$$

$$g_0(M,z) \;=\; 0. \tag{5.14}$$

Solving the second order difference equation (5.12) with boundary conditions (5.13) and (5.14) and substituting the values of p and q in (3.10), we obtain $g_0(N,z)$ as

$$g_0(N,z) \;=\; \frac{(2z)^N \{\, \theta_1(z)^{M-N} - \theta_2(z)^{M-N} \,\}}{\{\, \theta_1(z)^M - \theta_2(z)^M \,\}}, \tag{5.15}$$

where

$$\theta_1(z) \;=\; 1+\rho + \{\, (1+\rho)^2 - 4\rho z^2 \,\}^{1/2} \tag{5.16}$$

$$\theta_2(z) \;=\; 1+\rho - \{\, (1+\rho)^2 - 4\rho z^2 \,\}^{1/2}. \tag{5.17}$$

However, from relation (5.1) and (5.5), we have

$$G_0(N,z) \;=\; \frac{g_0(N,z)}{P(U_0)}. \tag{5.18}$$

Thus, substituting $g_0(N,z)$ in (5.15) into relation (5.18), we obtain $G_0(N,z)$, the probability generating functions of the number of transitions during U_0 as

$$G_0(N,z) = \frac{(1-\rho^M)(2z)^N\{\theta_1(z)^{M-N} - \theta_2(z)^{M-N}\}}{(1-\rho^{M-N})\{\theta_1(z)^M - \theta_2(z)^M\}}. \tag{5.19}$$

Similarly, we can obtain $G_1(N,z)$, the probability generating function of the number of transitions during U_1 as

$$G_1(N,z) = \frac{(1-\rho^M)(2\rho z)^N\{\theta_1(z)^N - \theta_2(z)^N\}}{(\rho^{M-N}-\rho^M)\{\theta_1(z)^M - \theta_2(z)^M\}}. \tag{5.20}$$

5.2.2 $G_0(K,z)$ and $G_1(K,z)$

Suppose that the number of customers in the system reaches K when both service stations are operating simultaneously. The number of customers in the system will fluctuate until it reaches 0 for the first time without reaching M with probability $P(U_2)$, or M for the first time without reaching 0 with probability $P(Y_1)$. However, the fluctuations of the number of customers in the system during U_2 and Y_1 are also dictated by the one step transition probability matrix of the number of customers in the system as shown in Figure 3.4. The values of p and q which are the elements of the one step transition probability matrix of J(n) are given in (3.10) because one service station is operating during U_2 and Y_1. Hence, $G_0(K,z)$ and $G_1(K,z)$ can be obtained using similar procedures when finding $G_0(N,z)$.

$G_0(K,z)$ and $G_1(K,z)$ can also be deduced directly from the results of $G_0(N,z)$ and $G_1(N,z)$ based on the same approaches used to find $P(U_2)$ and $P(Y_1)$ from $P(U_0)$ and $P(U_1)$ in Chapter 4. Thus, replacing N by K in

the results of $G_0(N,z)$ and $G_1(N,z)$ in (5.19) and (5.20), we obtain $G_0(K,z)$ and $G_1(K,z)$, the probability generating functions of the number of transitions during U_2 and Y_1 as

$$G_0(K,z) = \frac{(1-\rho^M)(2z)^K\{\ \theta_1(z)^{M-K} - \theta_2(z)^{M-K}\ \}}{(1-\rho^{M-K})\{\ \theta_1(z)^M - \theta_2(z)^M\ \}}, \tag{5.21}$$

$$G_1(K,z) = \frac{(1-\rho^M)(2\rho z)^K\{\ \theta_1(z)^K - \theta_2(z)^K\ \}}{(\rho^{M-K}-\rho^M)\{\ \theta_1(z)^M - \theta_2(z)^M\ \}}, \tag{5.22}$$

where $\theta_1(z)$ and $\theta_2(z)$ are also in (5.16) and (5.17).

5.2.3 $G_2(M,z)$

$G_2(M,z)$, the probability generating function of L_4, the number of transitions during V_r, can also be obtained by a similar method used to find $G_0(N,z)$. Let $P[L_4 = m]$ for $m=0,1,2,\cdots$ be the probability mass function of L_4. Then, $G_2(M,z)$ is defined by

$$G_2(M,z) = \sum_{m=0}^{\infty} z^m\ P[L_4 = m]. \tag{5.23}$$

After the number of customers in the system reaches M when one service station is operating, the level of customers in the system may fluctuate between K+1 and an indefinite level which is equal to or greater than M before it is decreased to K for the first time. Because the system operates under steady-state conditions, the probability that eventually the number of customers in the system reaches K during V_r after a finite number of transitions is one. However, the fluctuation of the number of customers in the system during V_r is dictated by the one step transition probability matrix of J(n) appearing in Figure 3.5. The values of p and q, which are the elements of the one step transition probability matrix of J(n) are shown in (3.11) because both service

stations are operating simultaneously during V_r. Hence, if we define

$$P_2(M,m) = P[L_4 = m] \quad \text{for all } m, \tag{5.24}$$

then, we have the following second order difference equation:

$$P_2(M,m+1) = p\,P_2(M+1,m) + q\,P_2(M-1,m) \quad \text{for } m \geq 0 \text{ and } K+1 \leq M \tag{5.25}$$

with boundary conditions

$$P_2(K,m) = 0 \qquad \text{for } m \geq 1 \tag{5.26}$$

$$P_2(K,0) = 1 \tag{5.27}$$

$$P_2(M,m) = 0 \qquad \text{for } m \geq 1. \tag{5.28}$$

Multiply (5.25) by z^{m+1} and add over all possible values of m. We obtain

$$G_2(M,z) = pz\,G_2(M+1,z) + qz\,G_2(M-1,z) \qquad \text{for } K < M \tag{5.29}$$

with boundary condition

$$G_2(K,z) = 1. \tag{5.30}$$

In order to solve completely the second order difference equation in (5.29), we need to specify another condition. The following indirect approach may be used to determine $G_2(M,z)$.

As shown in Figure 3.5, the one step transition probability matrix of the fluctuation of the number of customers in the system during V_r has an infinite number of states because the number of customers in the system can be increased to infinity. This situation eliminates one boundary condition of the difference equation (5.29). Hence, consider a problem in which we impose an upper restriction in the fluctuation of the number of customers in the system, say X where $X \leq M$. Then, the one step transition probability matrix for the fluctuation of the number of customers in the system will have a finite number of states. Thus, the

number of customers in the system can only fluctuate between K+1 and X-1 before it reaches K for the first time without reaching X, or X for the first time without reaching K. This implies that we can have two boundary conditions for the second order difference equation in (5.29). Let $g_2(M,z)$ be the solution of the following second order difference equation:

$$g_2(M,z) = pz\, g_2(M+1,z) + qz\, g_2(M-1,z) \qquad \text{for} \quad K < M \tag{5.31}$$

with boundary conditions

$$g_2(M,z) = 1 \tag{5.32}$$

$$g_2(X,z) = 0. \tag{5.33}$$

Suppose that we increase the value of X to infinity. Then, the number of customers in the system can fluctuate without any upper restrictions until it is absorbed to K. Therefore, as the value of X goes to infinity, $g_2(M,z)$ becomes $G_2(M,z)$ such that

$$G_2(M,z) = \lim_{X\to\infty} g_2(M,z). \tag{5.34}$$

Hence, we obtain the probability generating function of the number of transitions during V_1, $G_2(M,z)$ as

$$G_2(M,z) = \left[\frac{2+\rho - \{(2+\rho)^2 - 8\rho z^2\}^{1/2}}{2\rho z} \right]^{M-K}. \tag{5.35}$$

All these explicit forms of the probability generating functions, $G_0(N,z)$, $G_1(N,z)$, $G_0(K,z)$, $G_0(N,z)$ and $G_2(M,z)$ are the basic information necessary to derive the Laplace transforms of the probability density functions of U_0, U_1, U_2, Y_1 and V_1 discussed in the next section.

The results of the probability generating functions could be used to derive $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$. Let $E[L_0]$, $E[L_1]$, $E[L_2]$, $E[L_3]$ and $E[L_4]$ be the expected value of L_0, L_1, L_2, L_3 and L_4, respectively. Then, from relation (3.12) to (3.16), we have

$$E[U_0] = E[T_1]\, E[L_0], \tag{5.36}$$

$$E[U_1] = E[T_1]\, E[L_1], \tag{5.37}$$

$$E[U_2] = E[T_1]\, E[L_2], \tag{5.38}$$

$$E[Y_1] = E[T_1]\, E[L_3], \tag{5.39}$$

$$E[V_1] = E[T_1]\, E[L_4]. \tag{5.40}$$

$E[T_1]$ was obtained as

$$E[T_1] = \begin{cases} \dfrac{1}{\mu(1+\rho)} & \text{for } \; O = 1 \qquad (5.41) \\[2ex] \dfrac{1}{\mu(2+\rho)} & \text{for } \; O = 2. \qquad (5.42) \end{cases}$$

When computing $E[U_0]$, $E[U_1]$, $E[U_2]$ and $E[Y_1]$, $E[T_1]$ for the case $O = 1$ in (5.41) is considered, while when computing $E[V_1]$, $E[T_1]$ for the case $O = 2$ in (5.42) is considered.

$E[L_0]$, $E[L_1]$, $E[L_2]$, $E[L_3]$ and $E[L_4]$ are obtained respectively from $G_0(N,z)$, $G_1(N,z)$, $G_0(K,z)$, $G_1(K,z)$ and $G_2(M,z)$ by differentiating with respect to z and setting $z = 0$. We obtain as an example $E[L_0]$ for $E[U_0]$. The other expressions in (5.37) to (5.40) are obtained similarly. $E[L_0]$ is obtained as

$$E[L_0] = \frac{(1+\rho)[N(1+\rho^{M-N})(1-\rho^M) - 2M(\rho^{M-N}-\rho^M)]}{(1-\rho)(1-\rho^M)(1-\rho^{M-N})}. \tag{5.43}$$

Substituting $E[T_1]$ in (5.41) and the above $E[L_0]$ into relation (5.36)

yields

$$E[U_0] = \frac{N(1+\rho^{M-N})(1-\rho^M) - 2M(\rho^{M-N}-\rho^M)}{\mu(1-\rho)(1-\rho^M)(1-\rho^{M-N})}. \tag{5.44}$$

The explicit forms of $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$ will be used to develop alternative methodologies to find certain system characteristics in Chapter 7.

5.3 Laplace Transforms of the Probability Density Functions of U_0, U_1, U_2, Y_1 and V_1

The Laplace transforms of the probability density functions of U_0, U_1, U_2, Y_1 and V_1 may also be used for finding the values of $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$. More importantly however, they are used to find the Laplace transforms of the probability density functions of B, C, B_1 and B_2 explicitly.

Let

1) $f_{U_0}(x)$ for $0 \le x < \infty$; the probability density function of U_0,

2) $\bar{f}_{U_0}(s)$, $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$: the Laplace transforms of the probability density function of U_0, U_1, U_2, Y_1 and V_1, respectively, i.e., for example,

$$\bar{f}_{U_0}(s) = \int_0^\infty e^{-sx} f_{U_0}(x)\, dx. \tag{5.45}$$

In order to find $\bar{f}_{U_0}(s)$, $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$, we use the convolution properties for the sum of independent random variables and the results of $G_0(N,z)$, $G_1(N,z)$, $G_0(K,z)$, $G_1(K,z)$ and $G_2(M,z)$ obtained in (5.19) to (5.22) and (5.35).

5.3.1 $\bar{f}_{U_0}(s)$

Define the following:

i) $f_T(x)$ for $0 \le x < \infty$; the common probability density function of T_n, the length of time between two successive transitions during U_0 (Note that the T_n's are identically, independently and exponentially distributed with parameter $\mu(1+\rho)$, and independent of L_0.),

ii) $\bar{f}_T(s)$; the Laplace transform of $f_T(x)$ such as

$$\bar{f}_T(s) = \int_0^\infty e^{-sx} f_T(x)\, dx, \tag{5.46}$$

iii) $f_{U_0}(x|L_0=m)$ for all m ; the conditional probability density function of U_0 given that the number of transitions during U_0 is m.

Then, we have

$$\bar{f}_T(s) = \frac{\mu(1+\rho)}{s+\mu(1+\rho)}. \tag{5.47}$$

Also, using relation (3.12) yields

$$f_{U_0}(x|L_0=m) = f_T^{*(m)}(x) \qquad \text{for all m,} \tag{5.48}$$

where $f_T^{*(m)}(x)$ denotes the m-fold convolution of $f_T(x)$ with itself. Thus, $f_{U_0}(x)$ is obtained from

$$f_{U_0}(x) = \sum_{m=0}^{\infty} f_{U_0}(x|L_0=m)\ P[L_0 = m]. \tag{5.49}$$

Substituting relation (5.48) into relation (5.49) and then taking the Laplace transform on both sides, we obtain

$$\bar{f}_{U_0}(s) = \sum_{m=0}^{\infty} [\bar{f}_T(s)]^m\ P[L_0 = m]. \tag{5.50}$$

However, from relation (5.1), we have

$$\bar{f}_{U_0}(s) = G_0[N, \bar{f}_T(s)]. \tag{5.51}$$

Relation (5.51) implies that $\bar{f}_{U_0}(s)$ can be obtained by substituting $\bar{f}_T(s)$ in (5.47) instead for z in $G_0(N,z)$ in (5.32). Hence, $\bar{f}_{U_0}(s)$, the of the probability density function of U_0 is obtained as

$$\bar{f}_{U_0}(s) = \frac{(1-\rho^M)(2\mu)^N[\ \{\pi_1(s)\}^{M-N} - \{\pi_2(s)\}^{M-N}\ ]}{(1-\rho^{M-N})[\ \{\pi_1(s)\}^M - \{\pi_2(s)\}^M\ ]} \tag{5.52}$$

where

$$\pi_1(s) = s+\lambda+\mu + \{\ (s+\lambda+\mu)^2 - 4\lambda\mu\ \}^{1/2} \tag{5.53}$$

$$\pi_2(s) = s+\lambda+\mu - \{\ (s+\lambda+\mu)^2 - 4\lambda\mu\ \}^{1/2}. \tag{5.54}$$

5.3.2 $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$

Similarly, we can obtain $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$ explicitly. During U_1, U_2 and Y_1, one service station is operating. Hence, substituting $\bar{f}_T(s)$ in (5.47) instead for z in $G_1(N,z)$, $G_0(K,z)$ and $G_1(K,z)$ derived in (5.20) to (5.22), we obtain $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$ and $\bar{f}_{Y_1}(s)$, the Laplace transforms of the probability density function of U_1, U_2 and Y_1 as

$$\bar{f}_{U_1}(s) = \frac{(1-\rho^M)(2\lambda)^{M-N}[\ \{\pi_1(s)\}^N - \{\pi_2(s)\}^N\ ]}{(\rho^{M-N} - \rho^M)[\ \{\pi_1(s)\}^M - \{\pi_2(s)\}^M\ ]}, \tag{5.55}$$

$$\bar{f}_{U_2}(s) = \frac{(1-\rho^M)(2\mu)^K[\ \{\pi_1(s)\}^{M-K} - \{\pi_2(s)\}^{M-K}\ ]}{(1-\rho^{M-K})[\ \{\pi_1(s)\}^M - \{\pi_2(s)\}^M\ ]}, \tag{5.56}$$

$$\bar{f}_{Y_1}(s) = \frac{(1-\rho^M)(2\lambda)^{M-K}[\ \{\pi_1(s)\}^K - \{\pi_2(s)\}^K\ ]}{(\rho^{M-K} - \rho^M)[\ \{\pi_1(s)\}^M - \{\pi_2(s)\}^M\ ]}, \tag{5.57}$$

where $\pi_1(s)$ and $\pi_2(s)$ are in (5.53) and (5.54).

However, during V_1, both service stations are operating simultaneously. Thus, the length of times between two successive transitions are identically, independently and exponentially distributed with parameter $\mu(2+\rho)$. The Laplace transform of the common probability density function of the length of times between two successive transitions during V_1 becomes $\mu(2+\rho)/\{s+\mu(2+\rho)\}$. Substituting this value instead for z in $G_2(M,z)$ in (5.11), we obtain $\bar{f}_{V_1}(s)$, the Laplace transform of the probability density function of V_1 as

$$\bar{f}_{V_1}(s) = \left[\frac{s+\lambda+2\mu - \{ (s+\lambda+2\mu)^2 - 8\lambda\mu \}^{1/2}}{2\lambda} \right]^{M-K} . \tag{5.58}$$

Next, we will derive the Laplace transforms of the probability density functions of B, C, B_1 and B_2.

5.4 Laplace Transforms of the Probability Density Functions of B, C, B_1 and B_2

In the previous section, we obtained the basic components for deriving the Laplace transforms of the probability density functions of the busy period B, of the busy cycle C, of the length of time when one service station is operating and of the length of time when both service stations are operating simultaneously during the busy period, B_1 and B_2. We now proceed with the necessary definitions to derive these Laplace transforms.

Define the following:

i) $f_B(x)$, $f_{B_1}(x)$, $f_{B_2}(x)$ and $f_C(x)$ for $0 \le x < \infty$; the probability density function of B, B_1, B_2 and C, respectively,

ii) $\bar{f}_B(s)$, $\bar{f}_{B_1}(s)$, $\bar{f}_{B_2}(s)$ and $\bar{f}_C(s)$; the Laplace transforms of $f_B(x)$, $f_{B_1}(x)$, $f_{B_2}(x)$ and $f_C(x)$, respectively,

iii) $\bar{f}_I(s)$; the Laplace transform of the probability density function of the idle period,

iv) $f_B(x|F_a=r)$, $f_{B_1}(x|F_a=r)$ and $f_{B_2}(x|F_a=r)$ for all r ; the conditional probability density function of B, B_1 and B_2 given that F_a, the number of activations or removals of service stations during the busy period is r,

v) $f_{U_1}(x)$, $f_{U_2}(x)$, $f_{Y_1}(x)$, $f_V(x)$ and $f_{V_1}(x)$ for $0 \le x < \infty$; the probability density functions of the random variables U_1, U_2, Y_1, V and V_1, respectively,

vi) $\bar{f}_V(s)$; the Laplace transform of $f_V(x)$.

5.4.1 Derivation of $\bar{f}_B(s)$

We first derive $\bar{f}_B(s)$ based on relation (4.3) using the convolution properties of the sum of independent random variables. The result of the probability mass function of F_a, $P[F_a = r]$ obtained in (4.19) and the Laplace transforms of the probability density functions of U_0, U_1, U_2, Y_1 and V_1 obtained in (5.52) to (5.58) are required for proceeding to our objectives.

From relation (4.3), we have $f_B(x|F_a=r)$ for all r as

$$f_B(x|F_a=1) = f_{U_0}(x), \tag{5.59}$$

$$f_B(x|F_a=2) = f_{U_1}(x)*f_{V_1}(x)*f_{U_2}(x), \tag{5.60}$$

$$f_B(x|F_a=r) = f_{U_1}(x)*f_{V_1}{}^{*(r-1)}(x)*f_{Y_1}{}^{*(r-2)}(x)*f_{U_2}(x) \quad \text{for } r \ge 3, \tag{5.61}$$

where * denotes convolution operation and $f^{*(r)}(x)$ denotes the r-fold convolution of f(x) with itself.

Thus, we obtain $f_B(x)$ is from

$$f_B(x) = \sum_{r=1}^{\infty} f_B(x|F_a=r)\ P[F_a= r]. \tag{5.62}$$

Substituting $f_B(x|F_a=r)$ for all r in (5.59) to (5.61) into (5.62), and then taking the Laplace transform on both sides, we obtain

$$\bar{f}_B(s) = \bar{f}_{U_0}(s)P[F_a=1] + \bar{f}_{U_1}(s)\bar{f}_{U_2}(s)\bar{f}_{V_1}(s) \sum_{r=2}^{\infty} [\bar{f}_{V_1}(s)\bar{f}_{Y_1}(s)]^{r-2}\ P[F_a=r]. \tag{5.63}$$

Substituting $P[F_a=r]$ in (4.18) into (5.63) and simplifying yield $\bar{f}_B(s)$, the Laplace transform of the probability density function of the busy period as

$$\bar{f}_B(s) = \left[\frac{1-\rho^{M-N}}{1-\rho^{M}}\right]\bar{f}_{U_0}(s) + \frac{(1-\rho^{M-K})(\rho^{M-N} - \rho^{M})\bar{f}_{U_1}(s)\bar{f}_{U_2}(s)\bar{f}_{V_1}(s)}{(1-\rho^{M})[1-\rho^{M} - (\rho^{M-K}-\rho^{M})\bar{f}_{V_1}(s)\bar{f}_{Y_1}(s)]} \tag{5.64}$$

where

$$\bar{f}_{U_0}(s) = \frac{(1-\rho^{M})(2\mu)^{N}[\ \{\pi_1(s)\}^{M-N} - \{\pi_2(s)\}^{M-N}\ ]}{(1-\rho^{M-N})[\ \{\pi_1(s)\}^{M} - \{\pi_2(s)\}^{M}\ ]}, \tag{5.65}$$

$$\bar{f}_{U_1}(s) = \frac{(1-\rho^{M})(2\lambda)^{M-N}[\ \{\pi_1(s)\}^{N} - \{\pi_2(s)\}^{N}\ ]}{(\rho^{M-N}-\rho^{M})[\ \{\pi_1(s)\}^{M} - \{\pi_2(s)\}^{M}\ ]}, \tag{5.66}$$

$$\bar{f}_{U_2}(s) = \frac{(1-\rho^{M})(2\mu)^{K}[\ \{\pi_1(s)\}^{M-K} - \{\pi_2(s)\}^{M-K}\ ]}{(1-\rho^{M-K})[\ \{\pi_1(s)\}^{M} - \{\pi_2(s)\}^{M}\ ]}, \tag{5.67}$$

$$\bar{f}_{Y_1}(s) = \frac{(1-\rho^{M})(2\lambda)^{M-K}[\ \{\pi_1(s)\}^{K} - \{\pi_2(s)\}^{K}\ ]}{(\rho^{M-K}-\rho^{M})[\ \{\pi_1(s)\}^{M} - \{\pi_2(s)\}^{M}\ ]}, \tag{5.68}$$

$$\bar{f}_{V_1}(s) = \left[\frac{s+\lambda+2\mu - \{ (s+\lambda+2\mu)^2 - 8\lambda\mu \}^{1/2}}{2\lambda} \right]^{M-K}, \tag{5.69}$$

$$\pi_1(s) = s+\lambda+\mu + \{ (s+\lambda+\mu)^2 - 4\lambda\mu \}^{1/2}, \tag{5.70}$$

$$\pi_2(s) = s+\lambda+\mu - \{ (s+\lambda+\mu)^2 - 4\lambda\mu \}^{1/2}. \tag{5.71}$$

From the Laplace transform of the probability density function of the busy period obtained in (5.61) to (5.68), we can deduce the Laplace transforms of the probability density functions of the busy periods for the special cases of the M/M/2 system operating under the (0,K,N,M) policy.

The M/M/2 system operating under the (0,1,1,2) policy is equivalent to the ordinary M/M/2 system. Hence, setting K=N=1 and M=2 in relations (5.64) to (5.69) yields the Laplace transform of the probability density function of the busy period for the ordinary M/M/2 system as

$$\bar{f}_B(s) = \frac{2\mu}{s+\lambda + \{(s+\lambda+2\mu)^2 - 8\lambda\mu\}^{1/2}}, \tag{5.72}$$

thus, recovering Conolly's result (see Conolly[8]).

Another special case of our model is the $(0,K,N,\infty)$ policy which is equivalent to the M/M/1 system operating under the N policy, originated by Yadin and Naor[27]. Hence, taking the limit in relation in (5.64) to (5.69) as M goes to infinity yields the Laplace transform of the probability density function of the busy period for the M/M/1 system operating under the N policy as

$$\bar{f}_B(s) = \left[\frac{2\mu}{s+\lambda+\mu + \{(s+\lambda+\mu)^2 - 4\lambda\mu\}^{1/2}} \right]^N. \tag{5.73}$$

Since the ordinary M/M/1 system is a special case of the M/M/1 system operating under the N policy (N=1), the Laplace transform of the probability density function of the busy period for the ordinary M/M/1 system is immediately obtained from (5.73) (see Conolly[8] or Kleinrock[15]).

We thus deduce that the busy period for the M/M/1 system operating under the N policy is the sum of N busy periods of the ordinary M/M/1 system.

Next, we will derive $\bar{f}_C(s)$, $\bar{f}_{B_1}(s)$ and $\bar{f}_{B_2}(s)$, the Laplace transforms of the probability density functions of C, B_1 and B_2.

5.4.2 Derivation of $\bar{f}_C(s)$, $\bar{f}_{B_1}(s)$ and $\bar{f}_{B_2}(s)$

$\bar{f}_C(s)$, the Laplace transform of the probability density function of the busy cycle can be obtained by applying the convolution properties to relation (2.29). From relation (2.31), we have

$$\bar{f}_I(s) = \left[\frac{\lambda}{s+\lambda}\right]^N . \tag{5.74}$$

However, from relation (2.29), we have

$$\bar{f}_C(s) = \bar{f}_I(s)\,\bar{f}_B(s). \tag{5.75}$$

Thus, substituting $\bar{f}_I(s)$ in (5.74) and the result of $\bar{f}_B(s)$ in (5.64) to (5.71) into (5.75) yields $\bar{f}_C(s)$ explicitly.

Let $\bar{f}_{B_1}(s)$ and $\bar{f}_{B_2}(s)$ be the Laplace transforms of the probability density functions of the length of time when one service station is operating, B_1 and of the length of time when both service stations are

operating simultaneously, B_2 during the busy period. Using the same procedures to find $\bar{f}_B(s)$ based on relation (3.6) yields $\bar{f}_{B_1}(s)$ as

$$\bar{f}_{B_1}(s) = \left[\frac{1-\rho^{M-N}}{1-\rho^M}\right]\bar{f}_{U_0}(s) + \frac{(1-\rho^{M-K})(\rho^{M-N}-\rho^M)\bar{f}_{U_1}(s)\bar{f}_{U_2}(s)}{(1-\rho^M)[1-\rho^M - (\rho^{M-K}-\rho^M)\bar{f}_{Y_1}(s)]}. \tag{5.76}$$

However, the random variable V in relation (3.7) was defined as

$$P[V = 0] = 1. \tag{5.77}$$

Hence, we have

$$\bar{f}_V(s) = P[V = 0] = 1. \tag{5.78}$$

Therefore, using relation (3.7), we can obtain $\bar{f}_{B_2}(s)$ as

$$\bar{f}_{B_2}(s) = \frac{1-\rho^{M-N} - (\rho^{M-K}-\rho^{M-N})\bar{f}_{V_1}(s)}{1-\rho^M - (\rho^{M-K}-\rho^M)\bar{f}_{V_1}(s)}. \tag{5.79}$$

Thus, we can obtain $\bar{f}_{B_1}(s)$ and $\bar{f}_{B_2}(s)$ explicitly by substituting $\bar{f}_{U_0}(s)$, $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$ in (5.65) to (5.71) into relation (5.78) and (5.79).

Even though the Laplace transforms of the probability density functions of B, B_1 and B_2 can be obtained, it is difficult to invert them in general. However, the distribution function of B, B_1 and B_2 can be derived explicitly in terms of convolutions of the probability density functions of U_0, U_1, U_2, Y_1 and V_1 (see e.g. Appendix A). Hence, it is possible to investigate the behavior of the distribution of B, B_1 and B_2 for selected numerical values of λ, μ, K, N and M. When ρ=1, K=N=2 and M=3, the distribution function of B_1, B_2 and B for

different values of μ are shown in Figure 5.1, 5.2 and 5.3, respectively. In particular, the distribution function of B_2 has a concentration of mass with magnitude $P[F_a=1]$ at x=0. This is due to the fact that both service stations cannot be operating simultaneously when the number of activations or removals of service stations during the busy period is one.

In the next chapter, we will develop alternative methodologies to find the system characteristics using the results of Chapter 3, 4 and 5.

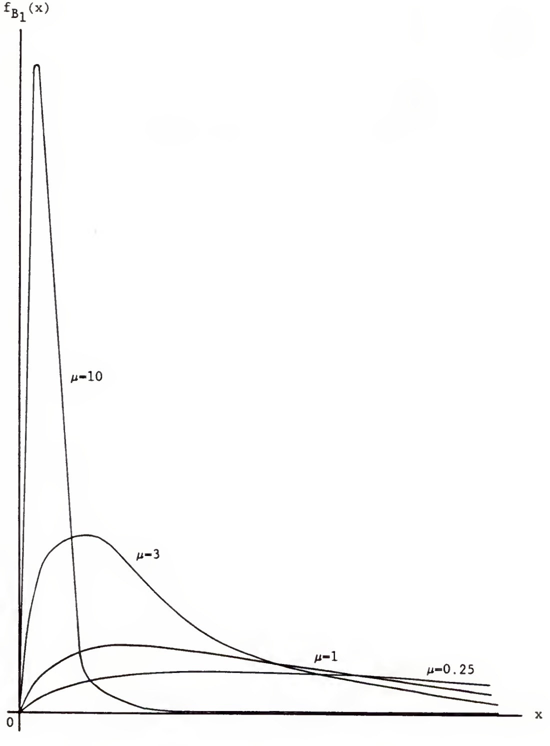

Figure 5.1 The probability density function of B_1 when μ = 10, 3, 1 and 0.25 given ρ = 1, K = N = 2 and M = 3

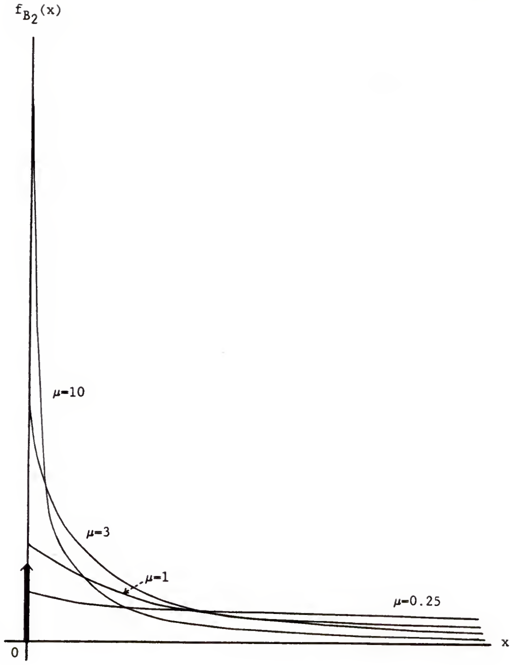

Figure 5.2 The mixed distribution function of B_2 when μ = 10, 3, 1 and 0.25 given ρ = 1, K = N = 2 and M = 3

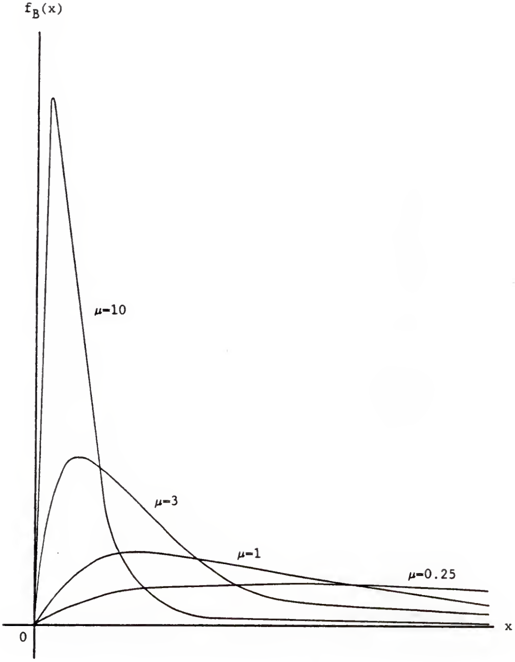

Figure 5.3 The probability density function of B when μ = 10, 3, 1 and 0.25 given ρ = 1, K = N = 2 and M = 3

CHAPTER 6

ALTERNATIVE METHODOLOGIES TO FIND SYSTEM CHARACTERISTICS

6.1 Introduction

The primary objective of this chapter is to derive alternative methodologies to find certain system characteristics that will be used to construct a total expected cost function in Chapter 7. Some of these characteristics were obtained in Chapter 2. However, the methodology does not use the explicit form of the steady-state joint probability mass function of the number of customers and the number of operating service stations in the system, P(i,j). It provides an alternative way to derive these system characteristics using the results of Chapter 3, 4 and 5.

The following is a list of the system characteristics to be obtained most of which being first moments:

(1) the expected number of activations or removals of service stations during the busy period, denoted by $E[F_a]$ or $E[F_r]$,

(2) the conditional expected length of times related to the busy period given that F_a is r for all r such as

1) the conditional expected length the busy period, denoted by $E[B|F_a=r]$,

2) the conditional expected length of the busy cycle, denoted by $E[C|F_a=r]$,

3) the conditional expected length of time when one service station is operating during the busy period, denoted by $E[B_1|F_a=r]$,

4) the conditional expected length of time when both service stations are operating simultaneously during the busy period, denoted by $E[B_2|F_a=r]$,

(3) the expected length of times related to the busy period such as

1) the expected length of the busy period, denoted by $E[B]$,

2) the expected length of the busy cycle, denoted by $E[C]$,

3) the expected length of time when i service stations are operating during the busy period, denoted by $E[B_i]$ for $i=1,2$,

(4) the steady-state probabilities such as

1) the probability that i service stations are operating during the busy cycle, denoted by $P[O=i]$ for $i=0,1,2$,

2) the steady-state probability that no service stations are operating and no customers are in the system, denoted by $P(0,0)$,

(5) the expected number of operating service stations, denoted by $E[O]$,

(6) the expected number of activations or removals of service stations per unit time, denoted by $E[R_a]$ or $E[R_r]$.

6.2 Expected Number of Activations or Removals of Service Stations During the Busy Period

The expected number of activations or removals of service stations during the busy period, $E[F_a]$ or $E[F_r]$ is obtained using the probability mass function of F_a, $P[F_a=r]$ obtained in (4.19) and (4.20). Hence, we have

$$E[F_a] = E[F_r] = \sum_{r=1}^{\infty} r\, P[F_a = r]$$

$$= \begin{cases} 1 + \dfrac{\rho^{M-N} - \rho^{M}}{1-\rho^{M-k}} & \text{for } \rho \neq 1 \\[2ex] 1 + \dfrac{N}{M-K} & \text{for } \rho = 1. \end{cases} \tag{6.1}$$

However, from relations in (3.3) and (6.1), $E[D]$ is also obtained as

$$E[D] = \sum_{d=0}^{\infty} d\, P[D = d]$$

$$= \begin{cases} \dfrac{\rho^{M-N} - \rho^{M}}{1-\rho^{M-k}} & \text{for } \rho \neq 1 \\[2ex] \dfrac{N}{M-K} & \text{for } \rho = 1. \end{cases} \tag{6.2}$$

6.3 Conditional Expected Length of Times related to the Busy Period

6.3.1 $E[B|F_a=r]$ for all r

Taking expectation on both sides of relation (3.4) or (4.3), we have for the conditional expected length of busy period given that F_a is r

$$E[B|F_a=r] = \begin{cases} E[U_0] & \text{for } r=1 \\ E[U_1] + E[U_2] + (r-1)E[V_1] + (r-2)E[V_1] & \text{for } r\geq 2. \end{cases} \tag{6.3}$$

In order to evaluate the values of $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$, differentiate $\bar{f}_{U_0}(s)$, $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$ in

(5.52) and (5.55) to (5.57) with respect to s and set s to zero, we obtain respectively

$$E[U_0] = \begin{cases} \dfrac{N(1+\rho^{M-N})(1-\rho^M) - 2M(\rho^{M-N}-\rho^M)}{\mu(1-\rho)(1-\rho^M)(1-\rho^{M-N})} & \text{for } \rho \neq 1 \\[2ex] \dfrac{N(2M-N)}{6\mu} & \text{for } \rho = 1, \end{cases} \tag{6.4}$$

$$E[U_1] = \begin{cases} \dfrac{M(1+\rho^M)(1-\rho^N) - N(1+\rho^N)(1-\rho^M)}{\mu(1-\rho)(1-\rho^N)(1-\rho^M)} & \text{for } \rho \neq 1 \\[2ex] \dfrac{(M+N)(M-N)}{6\mu} & \text{for } \rho = 1, \end{cases} \tag{6.5}$$

$$E[U_2] = \begin{cases} \dfrac{K(1+\rho^{M-K})(1-\rho^M) - 2M(\rho^{M-K}-\rho^M)}{\mu(1-\rho)(1-\rho^M)(1-\rho^{M-K})} & \text{for } \rho \neq 1 \\[2ex] \dfrac{K(2M-K)}{6\mu} & \text{for } \rho = 1, \end{cases} \tag{6.6}$$

$$E[Y_1] = \begin{cases} \dfrac{M(1+\rho^M)(1-\rho^K) - K(1+\rho^K)(1-\rho^M)}{\mu(1-\rho)(1-\rho^K)(1-\rho^M)} & \text{for } \rho \neq 1 \\[2ex] \dfrac{(M+K)(M-K)}{6\mu} & \text{for } \rho = 1, \end{cases} \tag{6.7}$$

$$E[V_1] = \frac{M-K}{\mu(2-\rho)}. \tag{6.8}$$

Note that $E[U_2]$ and $E[Y_1]$ can also be obtained directly from the results of $E[U_0]$ and $E[U_1]$ by replacing N by K. Also note that relations (5.36) to (5.40) with $G_0(N,z)$, $G_1(N,z)$, $G_0(K,z)$, $G_1(K,z)$ and

$G_2(M,z)$ in (5.19) to (5.22) and (5.35) yields the same $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$ as in (6.4) to (6.8).

Substitute $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$ in (6.4) to (6.8) into relation (6.3) and simplify; we obtain

$$E[B|F_a{=}r] = \begin{cases} \dfrac{N(1+\rho^{M-N})(1-\rho^M) - 2M(\rho^{M-N}-\rho^M)}{\mu(1-\rho)(1-\rho^M)(1-\rho^{M-N})} & \text{for } = 1 \text{ and } \rho \neq 1 \\[2ex] \dfrac{N(2M-N)}{6\mu} & \text{for } r = 1 \text{ and } \rho = 1 \\[2ex] \dfrac{1}{\mu(1-\rho)}\left[\dfrac{2K(1-\rho^M)}{(1-\rho^K)(1-\rho^{M-K})} - \dfrac{2M(\rho^{M-K}-\rho^M)}{(1-\rho^M)(1-\rho^{M-K})} - \dfrac{N(1+\rho^N)}{(1-\rho^N)}\right] & \\[2ex] \quad + \dfrac{(r-1)}{\mu(1-\rho)(2-\rho)}\left[\dfrac{M(3-2\rho+\rho^M)}{(1-\rho^M)} - \dfrac{K(3-2\rho+\rho^K)}{(1-\rho^K)}\right] & \text{for } r \geq 2 \text{ and } \rho \neq 1 \\[2ex] \dfrac{2MK - N^2 + (r-1)(M-K)(6+M+K)}{6\mu} & \text{for } r \geq 2 \text{ and } \rho = 1. \end{cases} \tag{6.9}$$

6.3.2 <u>$E[C|F_a{=}r]$ for all r</u>

The conditional expected length of the busy cycle given that F_a is r, $E[C|F_a{=}r]$ can be obtained from

$$E[C|F_a{=}r] = E[I] + E[B|F_a{=}r], \tag{6.10}$$

where $E[I]$ is the expected length of the idle period and is given by

$$E[I] = N/\lambda.$$

6.3.3 $E[B_1|F_a=r]$ for all r

Using a similar procedure for finding $E[B|F_a=r]$, the conditional expected length of time when exactly one service station is operating during the busy period given that F_a is r, $E[B_1|F_a=r]$ can be obtained.

From relation (3.6), we have

$$E[B_1|F_a=r] = \begin{cases} E[U_0] & \text{for } r = 1 \\ E[U_1] + E[U_2] + (r-2)E[V_1] & \text{for } r \geq 2. \end{cases} \tag{6.11}$$

Thus, substituting $E[U_0]$, $E[U_1]$, $E[U_2]$ and $E[Y_1]$ in (6.4) to (6.7) into relation (6.11) and simplifying yield

$$E[B_1|F_a=r] = \begin{cases} \dfrac{N(1+\rho^{M-N})(1-\rho^M) - 2M(\rho^{M-N}-\rho^M)}{\mu(1-\rho)(1-\rho^M)(1-\rho^{M-N})} & \text{for } r = 1 \text{ and } \rho \neq 1 \\[2ex] \dfrac{N(2M-N)}{6\mu} & \text{for } r = 1 \text{ and } \rho = 1 \\[2ex] \dfrac{1}{\mu(1-\rho)}\left[\dfrac{2K(1-\rho^M)}{(1-\rho^K)(1-\rho^{M-K})} - \dfrac{2M(\rho^{M-K}-\rho^M)}{(1-\rho^M)(1-\rho^{M-K})} - \dfrac{N(1+\rho^N)}{(1-\rho^N)}\right] & \\ \quad + \dfrac{(r-1)}{\mu(1-\rho)}\left[\dfrac{M(1+\rho^M)}{(1-\rho^M)} - \dfrac{K(1+\rho^K)}{(1-\rho^K)}\right] & \text{for } r \geq 2 \text{ and } \rho \neq 1 \\[2ex] \dfrac{2MK - N^2 + (r-1)(M-K)(M+K)}{6\mu} & \text{for } r \geq 2 \text{ and } \rho = 1. \end{cases} \tag{6.12}$$

6.3.4 $E[B_2|F_a=r]$ for all r

From relation (3.7), we obtain the conditional expected length of time when both service stations are operating simultaneously during the busy period given that F_a is r, $E[B_2|F_a=r]$ as

$$E[B_2|F_a=r] = \begin{cases} 0 & \text{for } r = 1 \\ (r-1)E[V_1] & \text{for } r \geq 2. \end{cases} \tag{6.13}$$

Substituting $E[V_1]$ in (6.8) into the above relation yields

$$E[B_2|F_a=r] = \begin{cases} 0 & \text{for } r = 1 \\ \dfrac{(r-1)(M-K)}{\mu(2-\rho)} & \text{for } r \geq 2. \end{cases} \tag{6.14}$$

Note that from relation (3.5), $E[B_2|F_a=r]$ can be obtained using the results of $E[B|F_a=r]$ and $E[B_2|F_a=r]$ from

$$E[B|F_a=r] = E[B_1|F_a=r] + E[B_2|F_a=r] \quad \text{for all r.} \tag{6.15}$$

6.4 Expected Length of Times related to the Busy Period

Using the explicit results of $P[F_a=r]$ obtained in (4.18) to (4.19), $E[B|F_a=r]$ in (6.8), $E[B_1|F_a=r]$ in (6.12) and $E[B_2|F_a=r]$ in (6.14), we can obtain the expected length of times E[B], E[C] and $E[B_i]$ for i=1,2 related to the busy period.

6.4.1 E[B]

The expected length of the busy period, E[B] can be obtained from

$$E[B] = \sum_{r=1}^{\infty} E[B|F_a=r]\, P[F_a=r]. \tag{6.16}$$

Hence, we have

$$E[B] = P[F_a=1]E[U_0] + \{1 - P[F_a=1]\}\{E[U_1] + E[U_2]\}$$
$$+ \{E[F_a] - 1\}E[V_1] + \{E[F_a] - 2 + P[F_a=1]\}E[Y_1], \qquad (6.17)$$

or equivalently

$$E[B] = P[D=0]E[U_0] + \{1 - P[D=0]\}\{E[U_1] + E[U_2]\}$$
$$+ E[D]E[V_1] + \{E[D] - 1 + P[D=0]\}E[Y_1]. \qquad (6.18)$$

Substituting $P[F_a = 1]$ and $E[F_a]$ in (4.18) to (4.19) and (6.1) into relation (6.17), or $P[D=1]$ and $E[D]$ in (4.20) to (4.21) and (6.2) into relation (6.18) yields

$$E[B] = \begin{cases} \dfrac{(1-\rho^{M-N})}{(1-\rho^{M})} E[U_0] + \dfrac{(\rho^{M-N}-\rho^{M})}{(1-\rho^{M})} \{E[U_1] + E[U_2]\} \\ \quad + \dfrac{(\rho^{M-N}-\rho^{M})}{(1-\rho^{M-K})} E[V_1] + \dfrac{(\rho^{M-N}-\rho^{M})(\rho^{M-K}-\rho^{M})}{(1-\rho^{M})(1-\rho^{M-K})} E[Y_1] & \text{for } \rho \neq 1 \\ \\ \dfrac{(M-N)}{M} E[U_0] + \dfrac{N}{M} \{E[U_1] + E[U_2]\} + \dfrac{N}{M-K} E[V_1] \\ \quad + \dfrac{N\,K}{M(M-K)} E[Y_1] \qquad \text{for } \rho = 1. \end{cases} \qquad (6.19)$$

Finally, substituting $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$ in (6.4) to (6.8) into relation (6.19) and simplifying, we obtain the same E[B] as derived in (2.34).

6.4.2 E[C]

From relation (2.30), we have

$$E[C] = N/\lambda + E[B]. \qquad (6.20)$$

Thus, substituting E[B] into the above relation yields the same E[C] as

in (2.35).

6.4.3 $E[B_i]$ for i=1,2

$E[B_1]$ and $E[B_2]$, the expected length of time when exactly one service station is operating and when both service stations are operating simultaneously during the busy period are obtained from

$$E[B_i] = \sum_{r=1}^{\infty} E[B_i | F_a=r]\, P[F_a=r] \quad \text{for } i=1,2. \tag{6.21}$$

Hence, we have

$$\begin{aligned} E[B_1] &= P[F_a=1]\, E[U_0] + \{1 - P[F_a=1]\}\{E[U_1] + E[U_2]\} \\ &\quad + \{E[F_a] - 2 + P[F_a=1]\}E[Y_1], \end{aligned} \tag{6.22}$$

$$E[B_2] = \{ E[F_a] - 1 \}\, E[V_1]. \tag{6.23}$$

Substituting $P[F_a=1]$ in (4.18) to (4.19), $E[F_a]$ in (6.1) and $E[U_0]$, $E[U_1]$, $E[U_2]$, $E[Y_1]$ and $E[V_1]$ in (6.4) to (6.8) into relations in (6.22) and (6.23) and simplifying, we obtain $E[B_1]$ and $E[B_2]$ as

$$E[B_1] = \begin{cases} \dfrac{1}{\lambda}\left[\dfrac{N\rho(1-\rho^{M-K}) - \rho^{M-N+1}(1-\rho^N)(M-K)}{(1-\rho)(1-\rho^{M-K})} \right] & \text{for } \rho \neq 1 \\[2ex] \dfrac{N(M+K-N)}{2} & \text{for } \rho = 1, \end{cases} \tag{6.24}$$

$$E[B_2] = \begin{cases} \dfrac{\rho^{M-N+1}(1-\rho^N)(M-K)}{\lambda(2-\rho)(1-\rho^{M-K})} & \text{for } \rho \neq 1 \\[2ex] \dfrac{N}{\lambda} & \text{for } \rho = 1. \end{cases} \tag{6.25}$$

6.5 Steady-State Probabilities

Without solving all of the steady-state difference equations in (2.1) to (2.10), it is possible to obtain:

1) the steady-state probability that i service stations are operating during the busy cycle, denoted by P[O=i] for i=0,1,2,
2) the steady-state probability that no service stations are operating and no customers are in the system, denoted by P(0,0),

using the explicit forms of E[I], $E[B_1]$, $E[B_2]$ and E[C].

6.5.1 P[O=i] for i=0,1,2

The probability that i service stations are operating during the busy cycle, P[O=i] can be expressed by the ratio between the expected length of time when i service stations are operating and the expected length of the busy cycle such as

$$P[O=0] = \frac{E[I]}{E[C]}, \tag{6.26}$$

$$P[O=1] = \frac{E[B_1]}{E[C]}, \tag{6.27}$$

$$P[O=2] = \frac{E[B_2]}{E[C]}. \tag{6.28}$$

Thus, substituting E[I] and E[C] into (6.26) and simplifying, we obtain the same P[O=0] as in (2.33). However, substituting $E[B_1]$, $E[B_2]$ and E[C] in (6.24), (6.25) and (2.35) into relation (6.27) and (6.28) yields

$$P[O=1] = \begin{cases} \dfrac{(2-\rho)[N(1-\rho^{M-K}) - \rho^{M-N+1}(1-\rho^{N})(M-K)]}{N(2-\rho)(1-\rho^{M-K}) - \rho^{M-N+1}(1-\rho^{N})(M-K)} & \text{for } \rho \neq 1 \\ \dfrac{M+K-N}{M+K-N+4} & \text{for } \rho = 1, \end{cases} \tag{6.29}$$

$$P[O=2] = \begin{cases} \dfrac{(1-\rho)\rho^{M-N+1}(1-\rho^{N})(M-K)}{N(2-\rho)(1-\rho^{M-K}) - \rho^{M-N+1}(1-\rho^{N})(M-K)} & \text{for } \rho \neq 1 \\ \dfrac{2}{M+K-N+4} & \text{for } \rho = 1. \end{cases} \tag{6.30}$$

6.5.2 P(0,0)

From the rate diagram of the system appearing in Figure 2.1, it is easy to verify that

$$P[O = 0] = N\, P(0,0). \tag{6.31}$$

Hence,

$$P(0,0) = \frac{P[O = 0]}{N}. \tag{6.32}$$

Substituting P[O=0] in (6.26) into relation (6.32) yields

$$P(0,0) = \frac{1}{\lambda E[C]}. \tag{6.33}$$

Thus, it is easy to verify that we have the same P(0,0) as in (2.21) if we substitute E[C] into relation (6.33).

6.6 Expected Number of Operating Service Stations

Since the obtained $P[O=i]$ $(i=0,1,2)$ is the probability mass function of the number of operating service stations, the expected number of operating service stations, $E[O]$ is obtained from

$$E[O] = \sum_{i=0}^{2} i\, P[O=i]. \tag{6.34}$$

Substituting $P[O=i]$ in (6.26) to (6.28) into relation (6.34), we obtain

$$E[O] = \frac{E[B_1]}{E[C]} + \frac{2\, E[B_2]}{E[C]}. \tag{6.35}$$

Hence, substituting $E[B_1]$, $E[B_2]$ and $E[C]$ in (6.24), (6.25) and (2.35) into relation (6.35) and simplifying, we obtain the same $E[O]$ as in (2.25). In particular, relation (6.35) will be used to construct a total expected cost function in Chapter 7.

6.7 Expected Number of Activations or Removals of Service stations per Unit Time

$E[R_a]$ or $E[R_r]$, the expected number of activations or removals of service stations per unit time can be obtained from the following five methods which will lead the same $E[R_a]$ as in (2.38) depending on the available information.

Method 1 : Using $E[F_a]$ and $E[C]$, we have

$$E[R_a] = E[R_r] = \frac{E[F_a]}{E[C]}. \tag{6.36}$$

Method 2 : Substituting relation (3.3) into (6.36) yields

$$E[R_a] = \frac{E[D]}{E[C]} + \frac{1}{E[C]}. \tag{6.37}$$

Note that relation (6.37) will be used to construct a total expected cost function per unit time in Chapter 7.

Method 3 : From relation (3.3) and (6.23), we have

$$E[D] = \frac{E[B_2]}{E[V_1]}. \tag{6.38}$$

From relation (6.28), we also have

$$E[B_2] = E[C]\ P[O = 2]. \tag{6.39}$$

Therefore, substituting $E[B_2]$ in (6.23) into relation (6.38) yields

$$E[D] = \frac{E[C]\ P[O = 2]}{E[V_1]}. \tag{6.40}$$

Substitute E[D] in (6.40) into relation (6.37), we obtain

$$E[R_a] = \frac{P[O = 2]}{E[V_1]} + \frac{1}{E[C]}. \tag{6.41}$$

Method 4 : Using relation (6.22), (6.27) and (6.37), we have

$$E[R_a] = \frac{P[O=0]}{E[Y_1]} + \frac{1}{E[C]} - \frac{P[F_a=1]E[U_0]}{E[C]E[Y_1]} - \frac{(1 - P[F_a=1])(E[U_1] + E[U_2] - E[Y_1])}{E[C]E[Y_1]}. \tag{6.42}$$

Method 5 : From relation (6.18), we have E[D] as

$$E[D] = \frac{E[B]-E[U_1]+E[Y_1]-E[U_2]+P[D=0](E[U_1]+E[U_2]-E[U_0]-E[Y_1])}{E[Y_1]+E[V_1]}. \tag{6.43}$$

Hence, substituting E[D] in (6.43) into relation (6.37) yields

$$E[R_a] = \frac{E[B]-E[U_1]+2E[Y_1]+E[V_1]-E[U_2]+P[D=0](E[U_1]+E[U_2]-E[U_0]-E[Y_1])}{E[C](E[Y_1]+E[V_1])}. \tag{6.44}$$

In the next chapter, we will construct a total expected cost function using some of the system characteristics derived in this chapter and determine the optimal customer levels (K^*,N^*,M^*) numerically based on minimizing the total cost.

CHAPTER 7

ECONOMIC ANALYSIS OF THE M/M/2 SYSTEM OPERATING UNDER THE TRIADIC (0,K,N,M) POLICY

7.1 Introduction

In previous chapters, we derived several important system characteristics for predicting system behavior. The problem that we address now consists in determining the optimal customer levels (K^*,N^*,M^*) at which service stations should be adjusted based on an economic criterion. The common criterion for the determination of the optimal operating policy in the area of multi-removable queueing models is to minimize the total expected cost involved in system operation. Several types of cost function could be selected. Usually, the total expected cost involves as components the expected holding cost for waiting customers, the expected operating cost for operating service stations and the expected switching cost for activations and removals of service stations.

It is important to point out that most of the cost functions used in the area of multi-removable queueing models are constructed under two assumptions. The first is the assumption of homogeneous structure for each cost element. In other words, the same amount of the holding cost for each customer per unit time, the operating cost for each operating service station per unit time and the switching cost for each activation or removal of service stations, are assumed regardless of the system

states. The other involves no consideration of the cost for service stations removed from the system to perform auxiliary tasks.

In this chapter, we construct a total expected cost function per unit time which could be applicable to a variety of real situations. The cost for service stations removed from the system associated with performance of auxiliary task is included in the cost elements. We assume that auxiliary task is always available for service stations removed from the system. We also assume that the structures for some of cost elements vary depending on the system states. We do so by relaxing the assumptions of homogeneous structures for these cost elements.

Following the construction of the total expected cost function per unit time, we determine the optimal customer levels (K^*, N^*, M^*) numerically at which service stations should be adjusted in order to minimize the total expected cost function. Finally, we perform a sensitivity analysis.

7.2 Construction of the Total Expected Cost Function

In order to construct the total expected cost function per unit time, we assume that the cost elements consist of (1) the holding cost for waiting customers, (2) the operating cost for operating the service stations, (3) the activation cost and the removal cost for activating and removing the service stations and (4) the cost for service stations removed from the system associated with performance of auxiliary tasks. We also assume that the cost for service stations removed from the system associated with performance of auxiliary tasks involves (i) the cost to start performing auxiliary task when removed from the system, (ii) the cost to perform auxiliary task and (iii) the cost to stop

performing auxiliary task when reactivated. However, the structures of all cost elements except the holding cost vary depending on the number of operating service stations.

Next, we define the following mathematical representations for the cost elements.

(a) $\underline{\alpha}$; the holding cost for one customer in the system per unit time : The holding cost can be considered as the penalty cost for delaying service to the customers who are waiting in the system.

(b) $\underline{\beta_i \ (i=1,2)}$; the operating cost for operating the service stations per unit time when i service stations are operating : The operating cost is incurred by the operating service stations to provide service for the customers.

(c) $\underline{\tau_a \text{ and } \tau_r}$; the activation cost and the removal cost for activating and removing the service stations, respectively : The activation cost is incurred each time when a service station starts operating while both service stations remain inoperative temporarily. The removal cost is incurred each time when the operating service station is removed from the system while the other service station is already removed from the system. In other words, the activation cost is incurred once at the beginning of the busy period and the removal cost is incurred once at the end of the busy period during the entire busy cycle.

(d) $\underline{\epsilon_s}$; the cost to start performing auxiliary task by the service stations : This cost is incurred each time when the service station removed from the system starts performing auxiliary task while the other service station is still operating. In other words, this cost is incurred whenever the number of customers in the system reaches K while

both service stations are operating simultaneously. Hence this cost is incurred D times during the busy cycle where D was defined in Chapter 3.

(e) δ_i (i=1,2) ; the cost to perform auxiliary task by the service stations per unit time when i service stations are performing auxiliary task

(f) ϵ_t ; the cost to stop performing auxiliary task by the service stations when reactivated : This cost is incurred each time when the service station removed from the system stops performing auxiliary task when reactivated while the other service station is already operating. In other words, this cost is incurred whenever the number of customers in the system reaches M while one service station is operating. Hence, this cost is incurred D times during the busy cycle.

(g) TC(K,N,M) ; the total expected cost function per unit time : This is the sum of (i) the expected holding cost per unit time, (ii) the expected operating cost per unit time, (iii) the expected activation cost and the removal cost per unit time and (iv) the expected cost per unit time incurred by the service stations removed from the system associated with their performance of auxiliary tasks.

Using the definitions of each cost element above and the system characteristics obtained in the previous chapters, the total expected cost function per unit time can be represented by

$$TC(K,N,M) = \alpha E[J] + \beta_1 \frac{E[B_1]}{E[C]} + \beta_2 \frac{E[B_2]}{E[C]} + (\tau_a + \tau_r)\frac{1}{E[C]}$$

$$+ \delta_1 \frac{E[B_1]}{E[C]} + \delta_2 \frac{E[I]}{E[C]} + (\epsilon_s + \epsilon_t)\frac{E[D]}{E[C]}, \tag{7.1}$$

where E[J] is the expected number of customers in the system,

$E[B_i]$ (i=1,2) is the expected length of time when i service stations are operating during the busy period,

E[C] is the expected length of the busy cycle,

E[I] is the expected length of the idle period,

E[D] is the expected number of times the level of customers in the system reaches M while one service station is operating during the busy period.

Let

$$\tau_{ar} = \tau_a + \tau_r, \tag{7.2}$$

$$\epsilon_{st} = \epsilon_s + \epsilon_t. \tag{7.3}$$

Then, substituting E[J], $E[B_1]$, $E[B_2]$, E[I], E[C] and E[D] obtained in (2.23) to (2.24), (6.24), (6.25), (2.32), (2.35) and(6.2) into relation in (7.1) and simplifying yields the total expected cost function per unit time explicitly as

$$TC(K,N,M) = \left[\begin{array}{l} \dfrac{\alpha\rho(3-2\rho)}{(1-\rho)(2-\rho)} + \beta_1 + \delta_1 + \dfrac{\alpha(M+k+1)}{2} \\[2ex] - \dfrac{\alpha N(1-\rho^{M-K})\{\ (M+K-N)(2-\rho) + 4\ \}}{2\{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^N)(M-K)\}} \\[2ex] + \dfrac{\lambda(1-\rho)(2-\rho)[\tau_{ar}(1-\rho^{M-K}) + \epsilon_{st}(\rho^{M-N}-\rho^M)]}{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^N)(M-K)} \\[2ex] + \dfrac{(1-\rho)(\beta_2-\beta_1-\delta_1)(M-K)\rho^{M-N+1}(1-\rho^N)}{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^N)(M-K)} \\[2ex] + \dfrac{(1-\rho)(\delta_2-\delta_1-\beta_1)N(1-\rho^{M-K})(2-\rho)}{N(1-\rho^{M-K})(2-\rho) - \rho^{M-N+1}(1-\rho^N)(M-K)} \quad \text{for } \rho \neq 1, \end{array} \right. \tag{7.4}$$

$$TC(K,N,M) = \left\{ \begin{array}{l} \alpha \left[\dfrac{M+3}{2} + \dfrac{6(N-1) + (K+N)(K-N) + M(N-4)}{3(M+K-N+4)} \right] \\ \\ + \dfrac{(\beta_1+\delta_1)(M+N-K) + 2(\beta_2+\delta_2)}{(M+K-N+4)} \\ \\ + \dfrac{2\lambda[\tau_{ar}(M-K) + \epsilon_{st}N]}{N(M-K)(M+K-N+4)} \qquad \text{for } \rho = 1. \end{array} \right. \tag{7.4}$$

In general, we assume that

$$\beta_1 \le \beta_2 \le 2\beta_1, \tag{7.5}$$

$$\delta_1 \le \delta_2 \le 2\delta_1. \tag{7.6}$$

The relations above are similar to the concepts of individual and joint ordering policies in inventory theory (see Sivazlian and Stanfel[21]). However, if we set

$$2\beta_1 = \beta_2 = 2\beta, \tag{7.7}$$

$$2\delta_1 = \delta_2 = 2\delta, \tag{7.8}$$

$$\tau_{ar} = \epsilon_{st} = \tau, \tag{7.9}$$

then, TC(K,N,M) in (7.1) becomes

$$TC(K,N,M) = \alpha E[J] + \beta E[O] + \delta(2-E[O]) + \tau E[R_a]. \tag{7.10}$$

TC(K,N,M) is a common form for the total expected cost function used in the area of the multi-removable service stations. Note that each cost element maintains the homogeneous structure. However, from relation (2.25), E[O] is not a function of the decision variables (K,N,M). Hence, for the determination of optimal customer levels which minimize TC(K,N,M), relation (7.10) is equivalent to

$$TC(K,N,M) = \alpha E[J] + \tau E[R_a]. \tag{7.11}$$

Next, we determine the optimal customer levels (K^*, N^*, M^*) numerically.

7.3 Determination of the Optimal Customer Levels

As shown in (7.4), it is very difficult to determine the optimal customer levels (K^*, N^*, M^*) analytically which minimize the total expected cost function per unit time because of the complexity of the optimization problem, i.e., a nonlinear integer minimization problem. One of the possible way to obtain (K^*, N^*, M^*) is a numerical approach.

We provide three numerical examples to obtain the optimal customer levels for given specific values of the system parameters, λ and μ and the cost elements, α, $\beta_i(i=1,2)$, $\delta_i(i=1,2)$, τ_{ar} and ϵ_{st}.

Example 1 : Determination of (K^*, N^*, M^*) given the value of ρ

Suppose $\rho = 0.6$. Also suppose α=\$1.0, β_1=\$3.0, β_2=\$5.0, δ_1=\$1.0, δ_2=\$1.5, τ_{ar}=\$10.0 and ϵ_{st}=\$5.0.

Table 7.1 exhibits the results of (K^*, N^*, M^*) and its corresponding $TC(K^*, N^*, M^*)$ for different values of λ and μ.

Table 7.1 (K^*, N^*, M^*) given $\rho = 0.6$

λ	μ	(K^*, N^*, M^*)	$TC(K^*, N^*, M^*)$
3	5	(3,5,10)	\$ 8.7319
6	10	(5,7,14)	10.8573
9	15	(7,8,18)	12.4544
15	25	(9,11,23)	14.9422
30	50	(13,15,32)	19.4979

Example 2 : Determination of M^* given K=N=1

Suppose $\lambda = 1$. Also suppose α=\$1.0, β_1=\$3.0, β_2=\$5.0, δ_1=\$1.0, δ_2=\$1.5, τ_{ar}=\$10.0 and ϵ_{st}=\$5.0.

Then, M^* is obtained for different values of μ and is shown in Table 7.2.

Table 7.2 M^* given K=M=1

λ	μ	M^*	TC(1,1,M^*)
1	4	13	\$ 9.0208
1	3	11	8.3889
1	2.5	10	7.9663
1	2.0	8	7.4883
1	1.5	6	7.0835
1	0.9	4	7.3003
1	0.7	3	7.9496
1	0.6	3	10.8437
1	0.5	2	15.6991

Example 3 : Determination of (K^*,N^*,M^*) given $2\beta_1=\beta_2$, $2\delta_2=\delta_2$, $\tau_{ar}=\epsilon_{st}$

Suppose $\rho = 0.6$ where $\lambda = 3.0$ and $\mu = 5.0$. Also suppose α=\$1.0, β_1=\$3.0, β_2=\$6.0, δ_1=\$1.0, δ_2=\$2.0 and $\tau_{ar}=\epsilon_{st}$=\$5.0.

Then, (K^*,N^*,M^*) and TC(K^*,N^*,M^*) are obtained as (1,1,9) and \$7.0573.

As shown in Table 7.1 and 7.2, it is easy to see that the optimal customer levels change significantly for different values of the system

parameters, λ and μ. As the values of λ and μ are increased, the optimal customer levels are also increased.

Next, we perform a sensitivity analysis.

7.4 Sensitivity Analysis

Based on the case when $\lambda = 3.0$ and $\mu = 5.0$ in Example 1, we perform a sensitivity analysis for changes in the optimal customer levels along with changes in one of the following cost elements:

(1) β_2,

(2) δ_2,

(3) τ_{ar},

(4) ϵ_{st}.

The values of the other cost elements remain the same as in Example 1.

The values of (K^*,N^*,M^*) and the corresponding $TC(K^*,N^*,M^*)$ are shown in Table 7.3 to 7.6. It is easy to see that the optimal customer levels do not change significantly for different values of a specific cost element. However, exact variations of (K^*,N^*,M^*) are not predictable.

In the next chapter, we will discuss topics for future research in the area of the controllable M/M/2 queueing system operating under the triadic (0,K,N,M) policy.

Table 7.3 Sensitivity analysis with changes in β_2

β_2	(K^*,N^*,M^*)	$TC(K^*,N^*,M^*)$
\$6.0	(3,5,11)	\$8.7455
5.5	(3,5,11)	8.7393
5.0	(3,5,10)	8.7319
4.5	(3,5,10)	8.7227
4.0	(3,5,10)	8.7134
3.5	(2,5,10)	8.7041
3.0	(2,5,10)	8.6932

Table 7.4 Sensitivity analysis with changes in δ_2

δ_2	(K^*,N^*,M^*)	$TC(K^*,N^*,M^*)$
\$1.0	(3,5,10)	\$8.5227
1.2	(3,5,10)	8.6064
1.4	(3,5,10)	8.6901
1.5	(3,5,10)	8.7319
1.7	(3,5,11)	8.8156
1.9	(3,5,11)	8.8982
2.0	(3,5,11)	8.9393

Table 7.5 Sensitivity analysis with changes in r_{ar}

r_{ar}	(K^*,N^*,M^*)	$TC(K^*,N^*,M^*)$
$10.0	(3,5,10)	$8.7319
9.0	(4,4,10)	8.4439
7.0	(3,3,9)	8.8059
5.0	(1,1,8)	6.8417
4.0	(1,1,7)	5.5942
2.0	(1,1,6)	3.0329
0.0	(1,1,5)	0.3737

Table 7.6 Sensitivity analysis with changes in ϵ_{st}

ϵ_{st}	(K^*,N^*,M^*)	$TC(K^*,N^*,M^*)$
$5.0	(3,5,10)	$8.7319
4.0	(3,5,10)	8.7319
3.0	(3,5,9)	8.7262
2.0	(3,5,8)	8.7111
1.0	(3,5,7)	8.6693
0.0	(4,5,6)	8.5789

CHAPTER 8

EXTENSIONS AND TOPICS FOR FUTURE RESEARCH

8.1 Introduction

Because the fluctuation of the number of customers in the system in a Markovian queueing system is a random walk, the properties of the gambler's ruin problem can provide fundamental information to characterize such system. Hence, the methodology used to analyze the busy period may be applicable to other Markovian queueing models.

In this chapter, we discuss the following in the area of the controllable M/M/2 queueing system operating under the triadic (0,K,N,M) policy as topics for future research:

(1) development of efficient algorithms or heuristic methods to obtain the optimal customer levels which minimize the total expected cost function per unit time derived in (7.4);

(2) analysis of the model with a finite system capacity, denoted by R;

(3) analysis of the model with a finite population size, denoted by Q.

Besides the above, construction of a total cost function based on a discounted cost criterion, analysis of the model with priority service discipline and analysis of the model with system dependent service rate are also suggested as topics for future research.

8.2 Development of Efficient Algorithms or Heuristic Methods to obtain the Optimal Customer Levels

As discussed in Chapter 7, it is very difficult to determine the optimal customer levels (K^*,N^*,M^*) analytically which minimize the total expected cost function TC(K,N,M) because of the complexity of the problem. Hence, one of the possible ways to obtain (K^*,N^*,M^*) is a numerical approach. However, the numerical approach necessitates a huge amount of computational effort because the values of TC(K,N,M) for all possible values of the decision variables (K,N,M) have to be evaluated. This situation leads towards the development of efficient algorithms or heuristic methods to obtain the optimal customer levels.

From relation (7.1), TC(K,N,M) is given by

$$TC(K,N,M) = \alpha E[J] + \beta_1\frac{E[B_1]}{E[C]} + \beta_2\frac{E[B_2]}{E[C]} + \tau_{ar}\frac{1}{E[C]}$$
$$+ \delta_1\frac{E[B_1]}{E[C]} + \delta_2\frac{E[I]}{E[C]} + \epsilon_{st}\frac{E[D]}{E[C]}. \tag{8.1}$$

It is clear that TC(K,N,M) depends of the values of the system parameters and the cost elements. Hence, an analysis of the functional behavior of E[J], $E[B_1]/E[C]$, $E[B_2]/E[C]$, $1/E[C]$, $E[I]/E[C]$ and $E[D]/E[C]$ might provide fundamental information to predict the behavior of TC(K,N,M). Also it might be necessary to analyze which cost elements are critical in evaluating TC(K,N,M) for different values of the system parameters. It may be possible to develop efficient algorithmic procedures or heuristic methods to obtain the optimal customer levels if additional analysis is performed in the structure of the cost function which takes the form of a nonlinear integer optimization problem.

8.3 Analysis of the Model with a Finite System Capacity

Suppose that the system capacity is limited to R which is a positive integer and is greater than M. Then, no customers are allowed to enter the system when R customers are in the system. As shown in the rate diagram of the model appearing in Figure 8.1, customers' mean arrival rate is still independent of the system states because of an infinite population size. The only difference between this model and the model with an infinite system capacity and an infinite population size analyzed in the previous chapters is the fluctuation of the number of customers in the system when both service stations are operating simultaneously. This implies that the results related to the fluctuation of the number of customers in the system when one service station is operating obtained in the previous chapters are directly applicable to this model. Hence, this model can be easily analyzed by deriving necessary information for the fluctuation of the number of customers in the system when both service stations are operating simultaneously.

The steady-state behavior of the model can be studied by solving the difference equations for the steady-state joint probability mass function $P(i,j)$ for $i=0,1,2$ and $j=0,1,2,\cdots,R$. The difference equations consist of (2.1) to (2.8), (2.9) and

$$(\lambda+2\mu)P(2,j) = 2\mu P(2,j+1) + \lambda P(2,j-1) \quad (K+2\le j\le R-1,\ R\ne M), \tag{8.2}$$

$$2\mu P(2,R) = \lambda P(2,R-1). \tag{8.3}$$

The values of p and q which are the elements of the one step transition probability matrix of the number of customers in the system

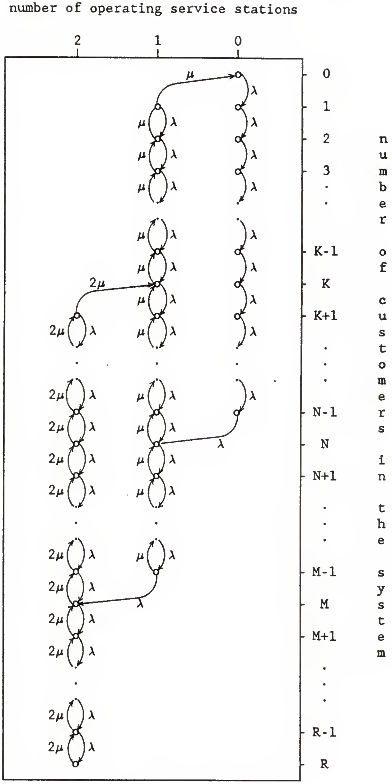

Figure 8.1 The rate diagram of the model with a finite system capacity

when both service stations are operating simultaneously remain the same as in (3.11). However, the one step transition probability that the number of customers in the system will reach R-1 after it reaches R is given by

$$P[J(n+1)=R-1|J(n)=R] = 1 \quad \text{for all } n. \tag{8.4}$$

Hence, the one step transition probability matrix of the number of the customers in the system when both service stations are operating simultaneously has a reflecting barrier at the state R as shown in Figure 8.2.

Under steady-state conditions, $P(V_1)$ which is the probability that the number of customers in the system will decreased to K after it reaches M when one service station is operating, is one. Therefore, the probability mass function of the number of activations or removals of service stations during the busy period is exactly the same as $P[F_a=r]$ obtained in (4.18) and (4.19).

In order to obtain the Laplace transform of the probability density function of the busy period, the probability generating function of the number of transitions during V_1, $G_2(M,z)$ has to be recomputed. This will lead to deriving the Laplace transform of the probability density function of V_1. $G_2(m,z)$ also satisfies a second order difference equation with boundary condition given in (5.29) and (5.30). Thus,

$$G_2(M,z) = pz\, G_2(M+1,z) + qz\, G_2(M-1,z) \quad \text{for } K<M<R-1 \tag{8.5}$$

with boundary condition

$$G_2(M,z) = 1. \tag{8.6}$$

However, another boundary condition related to $G_2(R,z)$ has to be evaluated based on the one step transition probability matrix of the

J(n)	K	K+1	K+2	⋯	M-1	M	M+1	⋯	R-2	R-1	R
K	1	0	0	⋯	0	0	0	⋯	0	0	0
K+1	q	0	p	⋯	0	0	0	⋯	0	0	0
K+2	0	q	0	⋯	0	0	0	⋯	0	0	0
.	.	.	.	⋯	.	.	.	⋯	.	.	.
M-1	0	0	0	⋯	0	p	0	⋯	0	0	0
M	0	0	0	⋯	q	0	p	⋯	0	0	0
M+1	0	0	0	⋯	0	q	0	⋯	0	0	0
.	.	.	.	⋯	.	.	.	⋯	.	.	.
R-2	0	0	0	⋯	0	0	0	⋯	p	0	0
R-1	0	0	0	⋯	0	0	0	⋯	q	0	p
R	0	0	0	⋯	0	0	0	⋯	0	1	0

Figure 8.2 The one step transition probability matrix of the number of customers in the system when both service stations are operating where the values of p and q are shown in (3.11)

number of customers in the system in which the state R forms a reflecting barrier. Thus, the difference equation (8.5) is solvable directly. Substituting $\{\mu(2+\rho)\}/\{s+\mu(2+\rho)\}$ instead for z in the solution $G_2(M,z)$ yields the Laplace transform of the probability density function of V_1. Therefore, substituting the Laplace transform of the probability density function of V_1 and the existing Laplace transforms of the probability density functions of U_0, U_1, U_2 and V_1 in (5.65) to (5.69) into relation (5.64) yields the Laplace transform of the probability density function of the busy period. Note that setting K=1, N=1 and M=2 yields the Laplace transform of the probability density function of the busy period for the ordinary M/M/2 queueing system with a finite system capacity R.

The total expected cost function per unit time can also be obtained by recomputing E[J], E[B_2] and E[C], and then substituting them into relation (7.1).

8.4 Analysis of the Model with a Finite Population Size

Suppose that the population size or the total number of possible customers is Q which is a positive integer and is greater than M. As shown in the rate diagram of the model appearing in Figure 8.3, customers' mean arrival rate depends on the number of customers in the system.

For the steady-state behavior of the model, the joint probability mass function P(i,j) for i=0,1,2 and j=0,1,2,$\cdots$,Q has to be recomputed using the following difference equations:

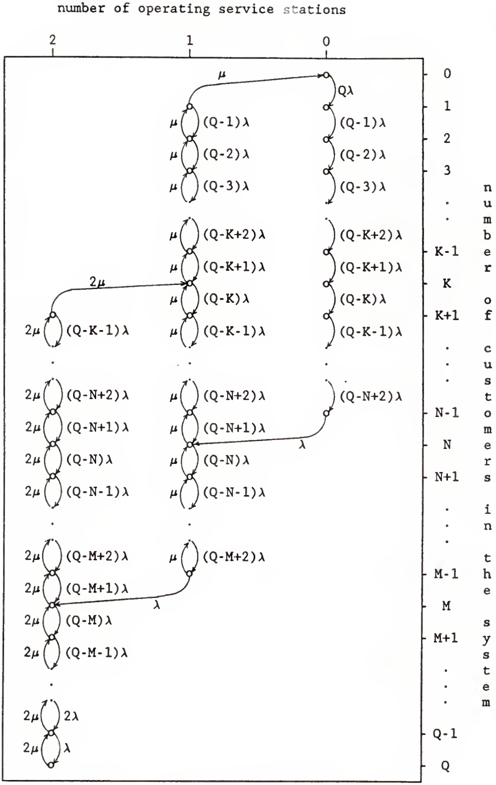

Figure 8.3 The rate diagram of the model with a finite population size

$$(Q-j)\lambda P(i,j) = (Q-j-1)\lambda P(i,j+1) \quad (i=0,\ 0\le j\le N-2), \tag{8.7}$$

$$Q\lambda P(i,j) = \mu P(i+1,j+1) \quad (i=0,\ j=0), \tag{8.8}$$

$$\{\mu+(Q-j)\lambda\}P(i,j) = \mu P(i,j+1) \quad (i=1,\ j=1), \tag{8.9}$$

$$\{\mu+(Q-j)\lambda\}P(i,j) = (Q-j+1)\lambda P(i,j-1) + \mu P(i,j+1) \quad (i=1,\ 2\le j\le M-2,\ j\ne K,M), \tag{8.10}$$

$$\{\mu+(Q-j)\lambda\}P(i,j) = (Q-j+1)\lambda P(i,j-1) + \mu P(i,j+1) + 2\mu P(i+1,j+1) \quad (i=1,j=K), \tag{8.11}$$

$$\{\mu+(Q-j)\lambda\}P(i,j) = (Q-j+1)\lambda P(i,j-1) + \mu P(i,j+1) + (Q-j+1)\lambda P(i-1,j-1) \quad (i=1,j=N), \tag{8.12}$$

$$\{\mu+(Q-j)\lambda\}P(i,j) = (Q-j+1)\lambda P(i,j-1) \quad (i=1,j=M-1), \tag{8.13}$$

$$\{2\mu+(Q-j)\lambda\}P(i,j) = 2\mu P(i,j-1) \quad (i=2,j=K+1), \tag{8.14}$$

$$\{2\mu+(Q-j)\lambda\}P(i,j) = (Q-j+1)\lambda P(i,j-1) + 2\mu P(i,j-1) \quad (i=2,\ K+2\le j\le Q-1,\ j\ne M), \tag{8.15}$$

$$\{2\mu+(Q-j)\lambda\}P(i,j) = (Q-j+1)\lambda P(i,j-1) + (Q-j+1)\lambda P(i-1,j-1) + 2\mu P(i,j+1) \quad (i=2,j=M), \tag{8.16}$$

$$2\mu P(i,j) = \lambda P(i,j-1) \quad (i=2,j=Q). \tag{8.17}$$

In order to analyze the busy period, suppose that i (i=1,2) service stations are operating. Let T_{ij} for $j=1,2,\cdots,Q$ be the random variable which represents the length of time from the instant when the number of customers in the system reaches j to the instant when the next transition occurs. Then, the T_{ij}'s are independently and exponentially distributed. It can be verified that the parameters for the probability density functions of the T_{ij}'s are $(Q-j)\lambda + i\mu$. Let

$$p_{ik} = P[J(n+1)=k|J(n)=j] \quad \text{for all } n. \tag{8.18}$$

In order to characterize the fluctuation of the number of customers in the system, the following one step transition probabilities of the number of the customers in the system can be written down:

$$p_{j,j+1} = \begin{cases} \dfrac{(Q-j)\lambda}{(Q-j)\lambda + \mu} & \text{for } O=1 \text{ and } j=1,2,\cdots,Q-1, \quad (8.19) \\[2ex] \dfrac{(Q-j)\lambda}{(Q-j)\lambda + 2\mu} & \text{for } O=2 \text{ and } j=1,2,\cdots,Q-1, \quad (8.20) \end{cases}$$

$$p_{j,j-1} = \begin{cases} \dfrac{\mu}{(Q-j)\lambda + \mu} & \text{for } O=1 \text{ and } j=1,2,\cdots,Q, \quad (8.21) \\[2ex] \dfrac{2\mu}{(Q-j)\lambda + 2\mu} & \text{for } O=2 \text{ and } j=1,2,\cdots,Q. \quad (8.22) \end{cases}$$

The fluctuation of the number of customers in the system during U_0, U_1, U_2 and Y_1 are also dictated by the one step transition probability matrix appearing in Figure 3.4. However, the values of p and q have to be replaced by $p_{j,i+1}$ and $p_{j,j-1}$, respectively, in (8.19) and (8.21). The one step transition matrix of the number of customers in the system during V_1 can also be obtained using the one step transition probability matrix shown in Figure 8.1 by replacing p and q by $p_{j,j+1}$ and $p_{j,j-1}$ in (8.20) and (8.22) after changing R to Q.

The probability mass function of the number of activations or removals of service stations during the busy period can be obtained from relation (4.7) because of

$$P(V_1) = 1. \tag{8.23}$$

However, the values of $P(U_0)$, $P(U_1)$, $P(U_2)$ and $P(Y_1)$ have to be recomputed. For example, $P(U_0)$ is obtained by solving the following second order difference equation where $\Phi_0(N)$ is equal to $P(U_0)$:

$$\Phi_0(N) = p_{N,N+1}\,\Phi_0(N+1) + p_{N,N-1}\,\Phi_0(N-1) \qquad \text{for } 1\le N\le M-1 \tag{8.24}$$

with boundary conditions

$$\Phi_0(0) = 1 \tag{8.25}$$

$$\Phi_0(M) = 0. \tag{8.26}$$

The values of $p_{N,N+1}$ and $p_{N,N-1}$ are shown in (8.19) and (8.21).

The Laplace transform of the probability density function of the busy period is obtained from relation (5.63) as

$$\bar{f}_B(s) = \bar{f}_{U_0}(s)P[F_a=r] + \bar{f}_{U_1}(s)\bar{f}_{U_2}(s)\bar{f}_{V_1}(s)\sum_{r=2}^{\infty}[\bar{f}_{V_1}(s)\bar{f}_{Y_1}(s)]^{r-2}P[F_a=r]. \tag{8.27}$$

Then, setting K=1, N=1 and M=2 in $\bar{f}_B(s)$ yields the Laplace transform of the probability density function of the busy period for the ordinary M/M/2 queueing system with a finite population size Q. However, a new methodology has to be developed in order to derive the explicit forms of the Laplace transform of the probability density functions of U_0, U_1, U_2, Y_1 and V_1. Since the T_{ij}'s are not identically distributed, the methodology used in Chapter 5 to obtain $\bar{f}_{U_0}(s)$, $\bar{f}_{U_1}(s)$, $\bar{f}_{U_2}(s)$, $\bar{f}_{Y_1}(s)$ and $\bar{f}_{V_1}(s)$ is not applicable to this model.

APPENDIX A

DERIVATION OF THE DISTRIBUTION FUNCTION OF B, B_1 AND B_2

In this appendix, we provide some analytic results for the distribution function of B, B_1, and B_2 discussed in Subsection 5.4.2.

Inverting the probability generating functions of L_0, L_1, L_2, L_3 and L_4 obtained in (5.19) to (5.22) and (5.35) yields the probability mass functions of the number of transitions during U_0, U_1, U_2, Y_1 and V_1 for m=1,2,$\cdots$ as (see e.g. Feller[10])

$$P[L_0=m] = \frac{(1-\rho^M)2^m\rho^{(m-N)/2}}{(1-\rho^{M-N})M(1+\rho)^m}\sum_{j=1}^{M-1}\sin(j\pi/M)\sin(jN\pi/M)\cos^{m-1}(j\pi/M), \tag{A.1}$$

$$P[L_1=m] = \frac{(1-\rho^M)2^m\rho^{(m+M-N)/2}}{(\rho^{M-N}-\rho^M)M(1+\rho)^m}\sum_{j=1}^{M-1}\sin(j\pi/M)\sin\{j(M-N)\pi/M\}\cos^{m-1}(j\pi/M), \tag{A.2}$$

$$P[L_2=m] = \frac{(1-\rho^M)2^m\rho^{(m-K)/2}}{(1-\rho^{M-K})M(1+\rho)^m}\sum_{j=1}^{M-1}\sin(j\pi/M)\sin(jK\pi/M)\cos^{m-1}(j\pi/M), \tag{A.3}$$

$$P[L_3=m] = \frac{(1-\rho^M)2^m\rho^{(m+M-K)/2}}{(\rho^{M-K}-\rho^M)M(1+\rho)^m}\sum_{j=1}^{M-1}\sin(j\pi/M)\sin\{j(M-K)\pi/M\}\cos^{m-1}(j\pi/M), \tag{A.4}$$

$$P[L_4=m] = \frac{M-K}{m}\binom{m}{(m+M-K)/2}\frac{\rho^{(m-M+K)/2}2^{(m+M-K)/2}}{(2+\rho)^m}. \tag{A.5}$$

The length of times between two successive transitions during the busy period are identically, independently and exponentially distributed with parameter $(\lambda+\mu)$ when one service station is operating and with

parameter $(\lambda+2\mu)$ when both service stations are operating simultaneously. Hence, $f_{U_0}(x|L_0=m)$ has a gamma distribution with parameter $(\lambda+\mu)$ and of order m such that

$$f_{U_0}(x|L_0=m) = \frac{(\lambda+\mu)^m}{\Gamma(m)} x^{m-1}e^{-(\lambda+\mu)x} \qquad \text{for } 0\le x<\infty. \tag{A.6}$$

Substituting $P[L_0=m]$ and $f_{U_0}(x|L_0=m)$ derived in (A.1) and (A.6) into relation (5.49) which is given by

$$f_{U_0}(x) = \sum_{m=0}^{\infty} f_{U_0}(x|L_0=m)\ P[L_0=m], \tag{A.7}$$

we obtain $f_{U_0}(x)$ for $0\le x<\infty$ as

$$f_{U_0}(x) = \frac{e^{-(\lambda+\mu)x}(1-\rho^M)}{(1-\rho^{M-N})Mx} \sum_{m=1}^{\infty} \frac{\{2(\lambda+\mu)x\}^m \rho^{(m-N)/2}}{\Gamma(m)(1+\rho)^m} \sum_{j=1}^{M-1} \theta_0(j,m), \tag{A.8}$$

where

$$\theta_0(j,m) = \sin(j\pi/M)\sin(jN\pi/M)\cos^{m-1}(j\pi/M).$$

Similarly, we obtain

$$f_{U_1}(x) = \frac{e^{-(\lambda+\mu)x}(1-\rho^M)}{(\rho^{M-N}-\rho^M)Mx} \sum_{m=1}^{\infty} \frac{\{2(\lambda+\mu)x\}^m \rho^{(m+M-N)/2}}{\Gamma(m)(1+\rho)^m} \sum_{j=1}^{M-1} \theta_1(j,m) \tag{A.9}$$

where

$$\theta_1(j,m) = \sin(j\pi/M)\sin\{j(M-N)\pi/M\}\cos^{m-1}(j\pi/M),$$

$$f_{U_2}(x) = \frac{e^{-(\lambda+\mu)x}(1-\rho^M)}{(1-\rho^{M-K})Mx} \sum_{m=1}^{\infty} \frac{\{2(\lambda+\mu)x\}^m \rho^{(m-K)/2}}{\Gamma(m)(1+\rho)^m} \sum_{j=1}^{M-1} \theta_2(j,m) \tag{A.10}$$

where

$$\theta_2(j,m) = \sin(j\pi/M)\sin(jK\pi/M)\cos^{m-1}(j\pi/M),$$

$$f_{Y_1}(x) = \frac{e^{-(\lambda+\mu)x}(1-\rho^M)}{(\rho^{M-K}-\rho^M)Mx} \sum_{m=1}^{\infty} \frac{\{2(\lambda+\mu)x\}^m \rho^{(m+M-K)/2}}{\Gamma(m)(1+\rho)^m} \sum_{j=1}^{M-1} \Theta_3(j,m) \qquad (A.11)$$

where

$$\Theta_3(j,m) = \sin(j\pi/M)\sin\{j(M-K)\pi/M\}\cos^{m-1}(j\pi/M),$$

$$f_{V_1}(x) = (M-K)e^{-(\lambda+2\mu)x} \sum_{m=0}^{\infty} \binom{m}{(m+M-K)/2} \frac{\rho^{(m-M+K)/2}2^{(m+M-K)/2}(\lambda+2\mu)^m x^{m-1}}{m\Gamma(m)(2+\rho)^m}. \qquad (A.12)$$

Using relations in (5.59) to (5.62) yield the probability density function of B, $f_B(x)$ for $0 \le x < \infty$ as

$$f_B(x) = f_{U_0}(x)P[F_a=1] + \sum_{r=2}^{\infty} f_{U_1}(x)*f_{V_1}{}^{*(r-1)}(x)*f_{Y_1}{}^{*(r-2)}(x)*f_{U_2}(x)\ P[F_a=r]. \qquad (A.13)$$

From relation (3.6), the probability density functions of B_1 for $0 \le x < \infty$ is represented by

$$f_{B_1}(x) = f_{U_0}(x)P[F_a=1] + \sum_{r=2}^{\infty} f_{U_1}(x)*f_{Y_1}{}^{*(r-2)}(x)*f_{U_2}(x)\ P[F_a=r], \qquad (A.14)$$

Since B_2 has a concentration of mass at x=0, we obtain from relation (3.7)

$$f_{B_2}(x) = \begin{cases} P[F_a=1] & \text{for } x=0 \\ \sum_{r=2}^{\infty} f_{V_1}{}^{*(r-1)}(x)\ P[F_a=r] & \text{for } 0<x<\infty. \end{cases} \qquad (A.15)$$

REFERENCES

[1] Balachandran, K.R. and Tijms, H., "On the D-Policy for the M/G/1 Queue," Management Science, Vol. 21, No. 9, 1073-1076, 1975.

[2] Bell, C.E., "Efficient Operation of Optimal-Priority Queueing System," Operations Research, Vol. 21, No. 3, 777-786, 1973.

[3] Bell, C.E., "Optimal Operation of an M/G/1 Priority Queue with Removable Server," Operations Research, Vol. 21, No. 6, 1281-1290, 1973.

[4] Bell, C.E., "Turning off a Server with Customers Present: Is this any way to run an M/M/c Queue with Removable Servers?," Operations Research, Vol. 23, No. 3, 571-573, 1975.

[5] Bell, C.E., "Optimal Operation of an M/M/2 Queue with Removable Servers," Operations Research, Vol. 28, No. 5, 1189-1204, 1980.

[6] Boxma, O.J., "Note on a Control Problem of Balachandran and Tijms," Management Science, Vol. 22, No. 8, 916-917, 1976.

[7] Cinlar, E., Introduction to Stochastic Processes, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

[8] Conolly, B., Lecture Notes on Queueing Systems, Halsted, New York, 1975.

[9] Crabill, T.B., Gross, D. and Magazine, M.J., " A Classified Bibliography of Research on Optimal Design and Control of Queues," Operations Research, Vol. 25, No. 2, 219-232, 1977.

[10] Feller, W., An Introduction to Probability Theory and its Applications, Vol. I, John Wiley & Sons, Inc., New York, 1968.

[11] Heyman, D., "Optimal Operating Policies for M/G/1 Queueing Systems," Operations Research, Vol. 16, No. 2, 362-382, 1968.

[12] Heyman, D., "The T-Policy for the M/G/1 Queue," Management Science, Vol. 23, No. 7, 775-778, 1977.

[13] Huang, C.C., Brunelle, S.L., Sawaki, K. and Vertinsky, I., "Optimal Control for Multi-Servers Queueing System under Periodic Review," Naval Research Logistics Quarterly, 127-135, 1977.

[14] Kimura, T., "Optimal Control of an M/G/1 Queueing System with Removable Server via Diffusion Approximation," European Journal of Operational Research, 390-398, 1981.

[15] Kleinrock, L., Queueing System Volume I: Theory, John Wiley & Sons, Inc., New York, 1975.

[16] Magazine, M.J., "Optimal Policies for Queueing System with Periodic Review," Ph.D. Dissertation, University of Florida, 1969.

[17] Magazine, M.J., "Optimal Control of Multi-Channel Service Systems," Naval Research Logistics Quarterly, 177-183, 1971.

[18] McGill, J.T., "Optimal Control of Queueing System with Variable Number of Enponential Servers," Technical Report No. 13, Dept. of Operations Research, Stanford University, August 24, 1969.

[19] Moder, J.J. and Phillips, Jr., C.R., "Queueing with Fixed and Variable Channels," Operations Research, Vol. 10, No. 2, 218-231, 1962.

[20] Prabhu, N.U., Queues and Inventories: A Study of their Basic Stochastic Processes, John Wiley & Sons, New York, 1965.

[21] Sivazlian, B.D. and Stanfel, L.E., Analysis of Systems in Operations Research, Prentice-Hall, Inc., New York, 1975.

[22] Sobel, M.J., "Optimal Average-Cost Policy for a Queue with Start-up and Shut-down Costs," Operations Research, Vol. 17, No. 1, 145-162, 1969.

[23] Szarkowicz, D.S. and Knowles, T.W., "Optimal Control of an M/M/s Queueing System," Operations Research, Vol. 33, No. 3, 644-660, 1985.

[24] Taylor, H.M. and Karlin, S., An Introduction to Stochastic Modeling, Academic Press, Inc., New York, 1984.

[25] Teghem, Jr., J., "Control of the Service Process in a Queueing System," European Journal of Operational Research, 141-158, 1986.

[26] Winston, W., "Optimal Monotonic Policies for Multiple-Server Exponential Queueing System with State Dependent Arrival Rates," Operations Research, Vol. 26, No. 26, 1089-1094, 1978.

[27] Yadin, M. and Naor, P., "Queueing System with Removable Service Stations," Operational Research Quarterly, Vol. 14, 393-405, 1963.

BIOGRAPHICAL SKETCH

Hahn-Kyou Rhee was born on December 25, 1954, in Andong, Korea. He was admitted at the HanYang University in Seoul, Korea, in March of 1974, where he received the degree of Bachelor of Industrial Engineering in February of 1978. From March, 1978 to February, 1983, he worked as a researcher in weapon systems R & D at the Agency for Defense Development (ADD) in Seoul, Korea. He was admitted to the Department of Industrial and Systems Engineering at the University of Florida in August of 1983 and received the degree of Master of Science with a major in industrial engineering in May of 1985. He continued his study there in pursuit of his doctoral degree in operations research.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Boghos D. Sivazlian, Chairman
Professor of Industrial and Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

J. G. Saw

John G. Saw
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Suleyman Tufekci
Associate Professor of Industrial and Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Chung-Yee Lee

Chung-Yee Lee
Assistant Professor of Industrial and Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sencer Yeralan

Sencer Yeralan
Assistant Professor of Industrial and Systems Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May, 1989

Hubert A. Bevis

Dean, College of Engineering

Dean, Graduate School